VYSOKÁ ŠKOLA BÁŇSKÁ – TECHNICKÁ UNIVERZITA OSTRAVA
EKONOMICKÁ FAKULTA

KATEDRA FINANCÍ

Využití analýzy přežití k predikci bankrotu

Application of Survival Analysis in Bankruptcy Prediction

Student:	Vojtěch Klimek
Vedoucí diplomové práce:	Ing. Novotná Martina, Ph.D.

Ostrava 2016

VŠB - Technická univerzita Ostrava
Ekonomická fakulta
Katedra financí

Zadání diplomové práce

Student: **Bc. Vojtěch Klimek**

Studijní program: N6202 Hospodářská politika a správa

Studijní obor: 6202T010 Finance

Téma: Využití analýzy přežití k predikci bankrotu
Application of Survival Analysis in Bankruptcy Prediction

Jazyk vypracování: čeština

Zásady pro vypracování:

1. Úvod
2. Charakteristika souhrnného hodnocení podniků
3. Teoretické vymezení analýzy přežití
4. Odhad modelů analýzy přežití
5. Závěr
Seznam použité literatury
Seznam zkratek
Prohlášení o využití výsledků diplomové práce
Seznam příloh
Přílohy

Seznam doporučené odborné literatury:

CLEVES, Mario A. *An introduction to survival analysis using Stata*. 3rd ed. College Station, Tex.: Stata Press, 2010. 412 s. ISBN 978-1-59718-074-0.
DLUHOŠOVÁ, Dana a kol. *Finanční řízení a rozhodování podniku: analýza, investování, oceňování, riziko, flexibilita*. 3. vyd. Praha: Ekopress, 2010. 225 s. ISBN 978-80-86929-68-2.
HOSMER, D., S. LEMESHOW and S. MAY. *Applied survival analysis: regression modeling of time-to-event data*. 2nd ed. Hoboken: Wiley-Interscience, 2008. 392 s. ISBN 978-0-471-75499-2.

Formální náležitosti a rozsah diplomové práce stanoví pokyny pro vypracování zveřejněné na webových stránkách fakulty.

Vedoucí diplomové práce: **Ing. Martina Novotná, Ph.D.**

Datum zadání: 20.11.2015
Datum odevzdání: 22.04.2016

Ing. Iveta Ratmanová, Ph.D.
vedoucí katedry

prof. Dr. Ing. Dana Dluhošová
děkanka fakulty

Prohlašuji, že jsem celou diplomovou práci, včetně všech příloh, vypracoval samostatně.

V Ostravě dne 21. 4. 2016

Klimek
....

Bc. Vojtěch Klimek

Poděkování

Tato diplomová práce vznikla za podpory SGS projektu VŠB-TU Ostrava 2015/75 Aplikace zobecněných lineárních modelů v pojišťovnictví a ve financích.

Na tomto místě bych také rád poděkoval vedoucí mé diplomové práce Ing. Martině Novotné, Ph.D. za odborné vedení, cenné rady a připomínky při konzultacích.

Obsah

1 Úvod

Zhodnocení finanční situace podniku je důležité pro mnoho různých subjektů. Ať už pro obchodní partnery dané společnosti, bankovní instituce, které se rozhodují o poskytnutí půjčky nebo pro podnik samotný. Z tohoto důvodu existuje velké množství metod, pomocí nichž je možné zhodnotit finanční situaci společnosti a neustále dochází k vyvíjení nových přístupů ve snaze o přesnější zachycení situace uvnitř podniku. Jedním z takových přístupů je analýza přežití, což je soubor statistických metod, který byl vyvinut zejména pro potřeby v medicíně, ale v současnosti je používán také v ekonomii a financích. Podstatou analýzy přežití je určit dobu přežití daného subjektu, zároveň je možné zahrnout i vliv nezávislých proměnných.

Cílem práce je posouzení doby přežití českých společností od jejich vzniku do bankrotu těchto podniků. Cíl je možné rozdělit na dvě části. V první části jsou porovnávány funkce přežití a hazardní funkce pro jednotlivá odvětví české ekonomiky. Analyzovány jsou podniky ze sektoru administrativa, informační technologie, kultura, stavebnictví, ubytování, zdravotnictví, zemědělství, zdravotní péče a společnosti zabývající se zpracováním vody a odpadu. Do analýzy vstupují firmy, které vznikly mezi roky 1990 až 2005, za předpokladu, že byly v roce 2005 stále fungující a byly u nich zveřejněny výsledky hospodaření za tento rok. Zdrojem pro tuto práci jsou jak podniky, u kterých bankrot nenastal, tak firmy, u kterých došlo k úpadku, přičemž k němu muselo dojít mezi roky 2005 až 2015. Druhým cílem této práce je sestavení Coxova semiparametrického modelu a parametrického modelu za předpokladu, že doba přežití odpovídá Weibullovu rozdělení pravděpodobnosti. Odhad modelů je proveden na základě údajů o českých společnostech a cílem odhadnutí těchto modelů je identifikovat takové finanční ukazatele, které ovlivňují dobu do bankrotu podniku a také vyčíslit vliv těchto ukazatelů.

Diplomová práce je rozčleněna do pěti kapitol, včetně úvodu a závěru. Po úvodu následují dvě teoretické kapitoly, obsahem čtvrté kapitoly je praktická část. V druhé kapitole je popsán bankrot společnosti a jeho možné příčiny, souvisejícími s riziky, kterým je společnost vystavena. Podstatnou částí této kapitoly je uvedení do problematiky bankrotních a bonitních modelů, sloužících k predikování výskytu bankrotu u podniku. Uveden je také popis matematicko-statistických metod, které jsou využívány pro odhad těchto modelů.

Prostřední část práce je zaměřena na detailní popis analýzy přežití. Nejprve jsou definovány základní pojmy v analýze přežití, jako je cenzorování, funkce přežití nebo hazardní funkce. Poté jsou charakterizovány jednotlivé metody, které jsou členěny na neparametrické, semiparametrické a parametrické. Na závěr kapitoly je vysvětlen princip sestavení modelu a jeho verifikace, a to jak statistická významnost odhadnutých parametrů, tak celková adekvátnost modelu.

Obsahem čtvrté kapitoly je aplikace metod, popsaných v kapitole třetí, na české společnosti. V první části jsou analyzovány podniky napříč různými sektory pomocí neparametrických metod, což umožňuje vzájemné srovnání jednotlivých odvětví. Dále jsou sestaveny dva modely pro sektor doprava, konkrétně semiparametrický Coxův model a parametrický model s proporcionálním hazardem a předpokladem, že doba přežití odpovídá Weibullovu rozdělení pravděpodobnosti. Tyto modely jsou verifikovány a na závěr je provedeno srovnání výsledků těchto modelů.

2 Charakteristika souhrnného hodnocení podniků

Obsahem kapitoly je definování bankrotu a popsání rizik, která mohou zapříčinit výskyt úpadku u společnosti. Další část je věnována způsobům, jak odhadnout, zda u podniku k bankrotu dojde nebo zda je společnost stabilní. Zdrojovými informacemi pro metody, jejichž cílem je odhadnout bankrot, je finanční analýza, která je rovněž popsána. Uveden je přehled komplexních modelů, sloužících k predikování bankrotu a popis matematicko-statistických metod, ze kterých tyto modely vycházejí.

2.1 Bankrot a jeho příčiny

U každého podniku může nastat situace, kdy nebude schopen plnit své závazky, což může vyústit v bankrot neboli úpadek společnosti. Obecně lze úpadek firmy charakterizovat jako ohrožení dalšího působení podniku kvůli nepříznivým jevům uvnitř podniku. U společnosti může nastat bankrot, pokud má dluhy, které není schopen splácet a hodnota těchto dluhů je vyšší než hodnota majetku daného podniku. Problematiku bankrotu upravuje v České republice Zákon o úpadku a způsobech jeho řešení, kde je definován dlužník v úpadku jako subjekt, který:

- má více věřitelů,
- má peněžité závazky po dobu delší než 30 dnů po době splatnosti a
- není schopen tyto závazky plnit, musí se přitom jednat o neschopnost, ne pouze o neochotu plnit závazky.

Přičemž je podle tohoto zákona uvažováno, že subjekt není schopen plnit své závazky pokud:

- zastavil platby podstatné části svých závazků, nebo
- je neplní po dobu delší než 3 měsíce po lhůtě splatnosti, nebo
- nesplnil povinnost předložit seznamy, které mu uložil insolvenční správce, nebo
- není možné pro věřitele dosáhnout uspokojení splatných peněžních pohledávek výkonem rozhodnutí nebo exekucí.

Pokud se u společnosti vyskytly známky bankrotu, pak můžou věřitelé podat insolvenční návrh. Po podání návrhu soud prozkoumá situaci u dlužníka a rozhodne, zda se jedná o bankrot, nebo ne. V případě, že se o bankrot jedná, soud rozhodne o způsobu řešení úpadku, přičemž existují čtyři možnosti. Jedná se o vyhlášení konkurzu, návrhu na reorganizaci společnosti, na oddlužení podniku anebo zvláštní způsob řešení, který stanoví Zákon o úpadku.

Důvodem, proč se finanční situace u podniku může dostat až do stavu, že na společnost je vyhlášen insolvenční návrh je skutečnost, že každý ekonomický subjekt je vystaven určitým rizikům. Riziko je možné podle Fotr (2005) chápat jako nebezpečí, že skutečné hospodářské výsledky podniku se budou lišit od předpokládaných výsledků. Pokud skutečné výsledky budou lepší než předpokládané, jedná se o pozitivní riziko, budou-li skutečné výsledky horší než předpokládané, jedná se o negativní riziko. Je samozřejmé, že problémem pro společnost je pouze riziko negativní. Riziko souvisí rovněž s pravděpodobností, neboť pro existenci rizika musí existovat pravděpodobnost, že nastane nějaká neočekávaná situace.

Rizika mohou být členěna podle mnoha hledisek, různí autoři uvádějí různá dělení, z nichž některá jsou nyní popsána. Základní rozdělení rizik je podle Smejkal (2006) na čisté a spekulativní riziko. Spekulativní riziko je využíváno ve financích a popisuje situaci, kdy může nastat kladná i záporná odchylka od očekávané hodnoty. Oproti tomu čisté riziko je používáno pro situaci, ve které je možná jenom ztráta a není možné dosáhnout pozitivních odchylek. Typickým čistým rizikem je ztráta na majetku v důsledku působení přírodních živlů.

Systematické riziko je podle Fotr (2005) takové riziko, kterým jsou v určité míře zasaženy všechny ekonomické subjekty v dané společnosti. Zdrojem tohoto rizika jsou například změny v daňových zákonech, fáze hospodářského cyklu, monetární a fiskální politika státu a podobně. Oproti tomu nesystematické riziko je jedinečné pro danou společnost, podle toho v jakém odvětví podniká a podle situace uvnitř firmy. Takovými riziky mohou pro podnik být vstup nového konkurenta na trh, špatné vedení firmy nebo nedostatečná inovativnost výrobků a služeb. Toto riziko je možné snížit diverzifikací, kdy firma bude podnikat ve více oblastech, bude mít více dodavatelů a odběratelů a sníží tak závislost na jednom produktu. Systematické riziko však snížit diverzifikací nelze, neboť jsou jim zasaženy veškeré subjekty na trhu.

S řízením rizik souvisí i členění rizik na ovlivnitelné a neovlivnitelné, což se týká toho, zda je vedení podniku schopné riziko eliminovat. Pokud lze snížit pravděpodobnost výskytu dané události, např. školením o bezpečnosti na pracovišti, pak se jedná o riziko ovlivnitelné. Pravděpodobnost výskytu neovlivnitelného rizika není možné snížit a patří tam veškerá systematická rizika. Management podniku může reagovat až na výskyt této události.

2.2 Finanční analýza

Významnou součástí finančního řízení podniku je finanční analýza. Jejím hlavní cílem je posouzení současné finanční situace a očekávané situace v budoucnosti. Finanční údaje o společnosti odrážejí úroveň podniku a jsou důležité pro spoustu subjektů, jako jsou vlastníci, banky, dodavatelé nebo management podniku.

Základními zdroji pro finanční analýzu jsou výkazy finančního účetnictví, které poskytují informace uživatelům mimo podnik. Tyto výkazy lze najít ve výročních zprávách firem a patří zde rozvaha, výkaz zisku a ztrát a obvykle taky výkaz cash flow. Pro řízení podniku jsou využívány výkazy vnitropodnikového účetnictví, které si podnik vytváří podle svých potřeb a nejsou přístupny veřejnosti. Zde se řadí nejrůznější členění nákladů, například druhové, kalkulační nebo podle jednotlivých středisek. Obecně lze údaje pro finanční analýzu rozčlenit dle Dluhošová (2010) na finanční informace (účetní výkazy, prognózy, burzovní informace), kvantifikovatelné nefinanční informace (statistiky produkce, odbytu, zaměstnanosti) a nekvantifikovatelné informace (zprávy vedoucích pracovníků, manažerů, odborného tisku).

Finanční analýza by měla komplexně zhodnotit finanční situaci podniku. Existuje proto spousta metod, jak analýzu provádět. První skupinou metod podle Dluhošová (2010) jsou deterministické metody, kde je pracováno s přesnými daty. Horizontální analýza je využívána k porovnání vývoje souhrnných hodnot v čase. Změna je vyčíslována jak v absolutních, tak v relativních hodnotách. Z historických údajů pak lze vyvodit dlouhodobý trend jednotlivých ukazatelů. Vertikální analýza oproti tomu slouží k posouzení významu jednotlivých složek ve vybraném souhrnném ukazateli včetně vývoje struktury v čase.

Pravděpodobně nejčastěji používanou a nejrychlejší metodou finanční analýzy je analýza poměrových ukazatelů. Jak napovídá název, poměrový ukazatel je dán jako poměr jedné položky z účetních výkazů k druhé. Výhodou těchto ukazatelů je, že převádí různé údaje na společnou úroveň, díky čemuž je poté jednoduché srovnání v čase nebo i s jinými podniky v odvětví. Poměrových ukazatelů je spousta a dělí se proto do různých skupin, kterými jsou ukazatele rentability, likvidity, aktivity a finanční stability a zadluženosti.

Druhou základní skupinou metod finanční analýzy jsou matematicko-statistické metody, kam patří například regresní analýza, diskriminační analýza, logistická regrese nebo analýza rozptylu. Podkladem pro tyto metody jsou často právě ukazatele z poměrové analýzy, proto jsou v další části tyto ukazatele popsány. Vzorce pro výpočet ukazatelů jsou z publikace Dluhošová (2010).

a) *Ukazatele rentability*

Základním cílem podnikání každého podniku by měl být zisk. Ukazatele rentability neboli výnosnosti ukazují, jak firma dokáže využít jednotku vloženého kapitálu. Rentabilita je tedy obecně určená jako poměr zisku a určité účetní položky. Podle typu zisku a typu položky je rozlišována rentabilita aktiv (ROA), rentabilita vlastního kapitálu (ROE), rentabilita tržeb (ROS) a rentabilita dlouhodobých zdrojů (ROCE). U všech těchto ukazatelů by měl být růstový trend v čase. Pro vypočítání ukazatelů rentability jsou využívány následující rovnice

$$\mathrm{ROA}=\frac{\mathrm{EBIT}}{\mathrm{A}}, \tag{2.1}$$

$$\mathrm{ROE}=\frac{\mathrm{EAT}}{\mathrm{VK}}, \tag{2.2}$$

$$\mathrm{ROS}=\frac{\mathrm{EAT}}{\mathrm{T}}, \tag{2.3}$$

$$\mathrm{ROC}=\frac{\mathrm{EAT}}{\text{náklady}}, \tag{2.4}$$

$$\mathrm{ROCE}=\frac{\mathrm{EBIT}}{\mathrm{VK+DCK}}, \tag{2.5}$$

kde EBIT představuje výsledek hospodaření před zdaněním a úroky, EAT je čistý zisk, A značí aktiva, VK je vlastní kapitál a DCK je dlouhodobý cizí kapitál.

b) *Ukazatele finanční stability a zadluženosti*

Ukazatele finanční stability a zadluženosti jsou používány při hodnocení, jak je podnik finančně zdravý a zda je stav jeho zadlužení udržitelný i do budoucnosti. Finanční stabilitu podniku je možné posoudit pomocí vztahu aktiv a pasiv. Při financování jednotlivých složek majetku cizími zdroji se hovoří o zadlužování, které může být v dané míře pro podnik přínosné, neboť může vést k vyšší rentabilitě. Vyšší zadlužování už však vede k riziku finanční nestability. Cílem každého podniku je tedy nalézt optimální poměr mezi cizími a vlastními zdroji. Mezi základní ukazatele lze zařadit následující.

Podíl vlastního kapitálu na aktivech (Equity Ratio) udává, do jaké míry je podnik schopen krýt svá aktiva vlastními zdroji a charakterizuje dlouhodobou finanční stabilitu. Výši ukazatele lze určit využitím vztahu

$$\text{Equity Ratio}=\frac{\text{vlastní kapitál}}{\text{aktiva}}. \tag{2.6}$$

Ukazatel celkové zadluženosti popisuje podíl celkových dluhů na aktivech. Čím vyšší je tento ukazatel, tím vyšší bude riziko pro věřitele, že podnik nesplatí své závazky. Vzorec platný pro tento ukazatel je

$$\text{Celková zadluženost} = \frac{\text{cizí kapitál}}{\text{aktiva}}. \tag{2.7}$$

Ukazatel zadluženosti vlastního kapitálu představuje míru finanční samostatnosti firmy a vyjadřuje výši dluhu připadající na jednu korunu vlastního kapitálu. Hodnota tohoto ukazatele by se měla pohybovat kolem 100 % a je možné jej vypočítat pomocí rovnice

$$\text{Zadluženost vlastního kapitálu} = \frac{\text{cizí kapitál}}{\text{vlastní kapitál}}. \tag{2.8}$$

Podíl dlouhodobého kapitálu na stálých aktivech je nazýván jako stupeň krytí stálých aktiv, ideální hodnota je 100 %, stálá aktiva by měla být krytá dlouhodobým kapitálem. Výpočet je zapisován jako

$$\text{Stupeň krytí stálých aktiv} = \frac{\text{dlouhodobý kapitál}}{\text{stálá aktiva}}. \tag{2.9}$$

Další ukazatel z této oblasti je finanční páka neboli majetkový koeficient, který je dán jako

$$\text{Majetkový koeficient} = \frac{\text{celková aktiva}}{\text{vlastní kapitál}}. \tag{2.10}$$

Poslední vybraný ukazatel z oblasti stability a zadluženosti je majetkový koeficient, který udává, kolikrát jsou nákladové úroky firmy kryty jejím ziskem, tedy

$$\text{Majetkový koeficient} = \frac{\text{EBIT}}{\text{úroky}}. \tag{2.11}$$

c) Ukazatele aktivity

Cílem výpočtu těchto ukazatelů je zjistit, jak efektivně dokáže podnik využívat svůj majetek. Využívají se dva typy ukazatelů - doba obratu a obrátka. Obrátkou je vyčíslováno, kolikrát za rok se daná položka přemění do jiné položky, nejčastěji do tržeb. Doba obratu naopak vyjadřuje délku období, které je nutné k uskutečnění jednoho obratu. Doba obratu je měřena u aktiv, zásob, pohledávek a závazků. Platí, že pokud se firma nechce dostat do finančních problémů, pak by doba obratu pohledávek měla být kratší než doba obratu závazků. Pro mezipodnikové srovnání je používána obrátka aktiv. Pro výpočet jsou vyžívány následující vztahy

$$\text{obrátka aktiv} = \frac{\text{tržby}}{\text{aktiva}}, \tag{2.12}$$

$$\text{doba obratu aktiv}=\frac{\text{aktiva}\cdot 360}{\text{tržby}}, \tag{2.13}$$

$$\text{doba obratu zásob}=\frac{\text{zásoby}\cdot 360}{\text{tržby}}, \tag{2.14}$$

$$\text{doba obratu pohledávek}=\frac{\text{pohledávky}\cdot 360}{\text{tržby}}, \tag{2.15}$$

$$\text{doba obratu závazků}=\frac{\text{závazky}\cdot 360}{\text{tržby}}, \tag{2.16}$$

d) Ukazatele likvidity

Kromě vysoké rentability a optimální výše zadlužení je podmínkou pro stabilní firmu rovněž likvidita neboli schopnost podniku dostát svým finančním závazkům. Ukazatele likvidity udávají, jak rychle dokáže podnik splatit své krátkodobé závazky. Rozlišuje se mezi ukazatelem celkové likvidity, pohotové likvidity a okamžité likvidity a to podle likvidnosti daného majetku. Pro správné řízení likvidity je dáno rozmezí hodnot, ve kterém by se ukazatele měly pohybovat. Výše ukazatele celkové likvidity by se měla pohybovat mezi 1,5 až 2,5, vztah je možné zapsat jako

$$\text{Celková likvidita}=\frac{\text{oběžná aktiva}}{\text{krátkodobé závazky}}. \tag{2.17}$$

U pohotové likvidity se počítá s likvidnějšími prostředky, neboť se od oběžných aktiv odečítají zásoby, které někdy v podnicích mohou být těžko směněny za peníze. Doporučená hodnota je 1 až 1,5, zápis rovnice je

$$\text{Pohotová likvidita}=\frac{\text{oběžná aktiva- zásoby}}{\text{krátkodobé závazky}}. \tag{2.18}$$

I v pohotové likviditě je ovšem zahrnutý majetek, který by v krátkém období mohl být obtížně převoditelný na peněžní prostředky. Využívá se proto i okamžitá likvidita, která bere v úvahu pouze pohotové platební prostředky, tedy peníze na účtu, v hotovosti a šeky. Ideální hodnota je mezi 0,2 až 0,5, výpočet tohoto ukazatele lze vyjádřit způsobem

$$\text{Okamžitá likvidita}=\frac{\text{pohotové platební prostředky}}{\text{krátkodobé závazky}}. \tag{2.19}$$

Vhodným doplňkem pro ukazatele likvidity jsou ukazatele struktury oběžných aktiv (OA), kde je rozlišován podíl pohledávek na oběžných aktivech a podíl zásob na oběžných aktivech, tedy

$$\text{Podíl pohledávek na OA}=\frac{\text{pohledávky}}{\text{oběžná aktiva}} \tag{2.20}$$

a

$$\text{Podíl zásob na OA}=\frac{\text{zásoby}}{\text{oběžná aktiva}}. \tag{2.21}$$

Důležitým ukazatelem je také ukazatel čistého pracovního kapitálu (ČPK), jehož výši určuje struktura aktiv a pasiv podniku. V ideálním případě by firma měla krýt stálá aktiva dlouhodobými zdroji a oběžná aktiva zdroji krátkodobými. Je více možností, jak tento ukazatel vypočítat, jednou z nich je

$$\text{ČPK}=\frac{\text{Oběžná aktiv}}{\text{krátkodobé závazky}}. \tag{2.22}$$

2.3 Bankrotní a bonitní modely

Pro svou jednoduchost a časovou nenáročnost jsou ke zjištění finanční situace využívány souhrnné modely, jejichž výsledkem je jedno číslo, které vystihuje situaci společnosti. Především v bankovnictví jsou často používány bankrotní a bonitní modely, které jsou postaveny na využití poměrových ukazatelů a přidělení určitých vah. Jak píše Dluhošová (2010), hlavním předpokladem pro použití těchto modelů je skutečnost, že v každém podniku mohou existovat znaky, pomocí nichž je možné predikovat zhoršení situace v budoucnu. Základní dělení je na modely bonitní a bankrotní. Cílem bonitních modelů je vyčíslit pravděpodobnost, že u podniku dojde ke zhoršení finanční úrovně. Do těchto modelů lze zařadit Tamariho model, Kralickuv Quick-test nebo index bonity. Úlohou bankrotních modelů je zhodnotit možnost, že u sledovaného podniku dojde k bankrotu. Typickými představiteli v této skupině je Altmanův model, Taflerův model, Beaverův model a Index IN. V rámci této podkapitoly jsou blíže představeny některé z těchto modelů.

2.3.1 Altmanův model

Jedním z prvních navrhnutých modelů byl Altmanův model, který se snaží pomocí hodnoty Z-skóre určit, zda u podniku dojde k bankrotu. Prvotní Altmanův model je určený pro veřejně obchodovatelné společnosti a obsahuje 5 poměrových ukazatelů. Hodnotu Z-skóre je možné určit podle Dluhošová (2010) jako

$$Z=3{,}3\cdot\frac{\text{Zisk}}{\text{Aktiva}}+\frac{\text{Tržby}}{\text{Aktiva}}+0{,}6\cdot\frac{\text{VK}}{\text{Celkové závazky}}+1{,}4\cdot\frac{\text{Nerozdělený zisk}}{\text{Aktiva}}+1{,}2\cdot\frac{\text{ČPK}}{\text{Aktiva}} \tag{2.23}$$

Podle výsledné hodnoty mohou být podniky zařazeny do třech kategorií. Pokud je hodnota Z-skóre vyšší než 2,99, pak je podnik důvěryhodný a pravděpodobnost bankrotu je jen minimální. Hodnota menší než 1,81 náleží firmám, u nichž je velké riziko bankrotu. Když je hodnota mezi těmito dvěma čísly, potom se jedná o šedou zónu a je u nich tedy určitá šance, že zbankrotují.

U tohoto modelu došlo k dalším modifikacím, konkrétně pro společnosti, které nejsou veřejně obchodovatelné a pro nevýrobní podniky. Někdy je také zmiňována a používaná česká varianta tohoto modelu, kdy je výsledná hodnota Z-skóre určena podle Vochozka (2011) vztahem

$$\begin{aligned} Z_{CZ} = {} & 3{,}3 \cdot \frac{\text{Zisk}}{\text{Aktiva}} + 0{,}99 \cdot \frac{\text{Tržby}}{\text{Aktiva}} + 0{,}6 \cdot \frac{\text{VK}}{\text{Celkové závazky}} + 1{,}4 \cdot \frac{\text{Nerozdělený zisk}}{\text{Aktiva}} \\ & + 6{,}56 \cdot \frac{\text{ČPK}}{\text{Aktiva}} - \frac{\text{Závazky po splatnosti}}{\text{Výnosy}} \end{aligned} \tag{2.24}$$

Modifikace spočívá ve zvýšení váhy pro ukazatel čistého pracovního kapitálu k aktivům a v přidání ukazatele poměru závazků po splatnosti k výnosům firmy. Hodnocení výsledné hodnoty je totožné jako u první varianty Altmanova modelu.

2.3.2 Index IN

Nejdůležitějším tuzemským bankrotním modelem, který zachycuje specifika českého trhu, je index podle Ivana a Inky Neumaierových. Ti zkoumali data 1000 českých firem a na jejich základě zkonstruovali první index IN95. Tento index byl dále upravován a obměňován a v současnosti existují jeho 4 varianty, přibyl index IN99, IN01 a IN05. Indexy se liší využitými poměrovými ukazateli a jejich váhami. K vypočtení je potřeba znalosti účetních dat, zejména se jedná o výši aktiv, cizích zdrojů, EBIT, nákladové úroky, celkové výnosy, oběžná aktiva, krátkodobý cizí kapitál a závazky po lhůtě splatnosti. Hodnotu zatím posledního indexu IN05 je možné zjistit využitím rovnice

$$\begin{aligned} \text{IN05} = {} & 0{,}13 \cdot \frac{\text{Aktiva}}{\text{Cizí zdroje}} + 0{,}04 \cdot \frac{\text{Zisk}}{\text{Nákl. úroky}} + 3{,}97 \cdot \frac{\text{Zisk}}{\text{Aktiva}} + 0{,}21 \cdot \frac{\text{Výnosy}}{\text{Aktiva}} \\ & + 0{,}09 \cdot \frac{\text{Oběžná aktiva}}{\text{Kr.cizí kapitál}} \end{aligned} \tag{2.25}$$

Výsledná hodnota tohoto indexu je opět zařazena do daného intervalu a podle toho je určeno, do jaké kategorie podnik spadá. Pokud je hodnota vyšší než 1,6, jedná se o bonitní podnik a riziko bankrotu je nízké. Hodnota menší než 0,9 naopak značí bankrotní podnik, pravděpodobnost bankrotu je vysoká. Pokud leží hodnota mezi 0,9 až 1,6, jde o šedou zónu a nelze určit, zda u podniku dojde k bankrotu, či nikoliv.

2.3.3 Kralickuv Quick-test

Zástupcem skupiny bonitních modelů je například Kralickuv Quick-test. Zde jsou sledovány čtyři poměrové ukazatele. Jedná se o podíl vlastního kapitálu na aktivech, podíl závazků na provozním cash flow, dále o poměr provozního cash flow a tržeb a rentabilitu aktiv. I u tohoto testu jsou podle Vochozka (2011) rozlišovány dvě varianty, původní a modifikovaná. V původní variantě je přiřazena známka pro každý ukazatel podle toho, do jakého intervalu spadá hodnota daného ukazatele pro konkrétní podnik. Oproti tomu v modifikované variantě jsou výsledné hodnoty poměřovány s percentily oborových hodnot. Tento postup je náročnější, neboť je potřeba zjistit percentily pro sledované odvětví. U obou variant existuje pětibodová stupnice, přičemž nejlepším hodnotám je přiřazena známka 1 a nejhorším známka 5.

Tímto testem je možné hodnotit výnosovou situaci podniku, kdy je zjištěna průměrná známka pro podíl vlastního kapitálu na aktivech a pro podíl závazků na provozním cash flow. Pokud je proveden aritmetický průměr známek pro poměr provozního cash flow a tržeb a pro rentabilitu aktiv, tak je ohodnocena finanční stabilita společnosti. Celkovou situaci firmy lze zjistit vypočtením aritmetického průměru známek pro výnosovou situaci a finanční stabilitu firmy. Přičemž pokud je známka nižší než 2, jedná se o bonitní podnik, známka mezi 2 a 3 značí šedou zónu a známka vyšší než 3 je typická pro bankrotní podniky. Rozdíl mezi tímto testem a předchozími bankrotními modely je, že výsledná hodnota je dána průměrem známek a tedy každý ukazatel má stejnou váhu.

2.3.4 Index bonity

Index bonity je dalším používaným bonitním modelem, a to především v evropských zemích. Index byl sestaven využitím diskriminační analýzy a k výpočtu výsledné hodnoty je nutné znát 6 účetních ukazatelů, ze kterých je sestaveno 6 poměrových ukazatelů, kterým jsou přiřazeny různé váhy a je možné tento vztah zapsat dle Vochozka (2011) jako

$$\begin{aligned} \text{IB}=0{,}08 \cdot \frac{\text{Aktiva}}{\text{Cizí zdroje}} + 1{,}5 \cdot \frac{\text{Cash flow}}{\text{Cizí zdroje}} + 10 \cdot \frac{\text{Zisk}}{\text{Aktiva}} + 5 \cdot \frac{\text{Zisk}}{\text{Výnosy}} \\ + 0{,}3 \cdot \frac{\text{Zásoby}}{\text{Výnosy}} + 0{,}1 \cdot \frac{\text{Výnosy}}{\text{Aktiva}} \end{aligned} \tag{2.26}$$

Hodnota indexu je nejvíce ovlivněna poměrem zisku a aktiv, tedy rentabilitou aktiv a také rentabilitou výnosu. Váhy pro ostatní ukazatele jsou výrazně nižší. Využitím tohoto modelu je možné určit nejen, zda je podnik bonitní, či bankrotní, ale také podrobnější ohodnocení finanční situace. Zásadním dělením je, zda je hodnota indexu (IB) kladná, nebo záporná. Pokud je hodnota záporná, pak je u společnosti velké riziko bankrotu. Oproti tomu kladná hodnota indexu značí bonitní firmu. Podrobnější hodnocení je znázorněno v tabulce (Tabulka 2.1).

Tabulka 2.1 Hodnocení podniku pomocí indexu bonity

Hodnota indexu bonity	(-∞; -2)	< -2; -1)	< -1; 0)	< 0; 1)	< 1; 2)	< 2; 3)	< 3; ∞)
Finanční situace podniku	extrémně špatná	velmi špatná	špatná	problematická	dobrá	velmi dobrá	extrémně dobrá

Zdroj: Vochozka (2011), vlastní zpracování

2.4 Matematicko-statistické metody

Výše uvedené bankrotní a bonitní modely byly odvozeny využitím matematicko-statistických metod. Tyto metody vycházejí z historických dat vybraného vzorku společností a jejich účelem je určit významné faktory, které ovlivňují finanční situaci podniku a odhadnout příslušné koeficienty. Výsledkem je model, do kterého jsou-li zadány data pro danou společnost, tak je možné zjistit, zda, nebo s jak velkou pravděpodobností, dojde k bankrotu firmy. Nejběžnějšími metodami pro tvorbu těchto modelů je vícenásobná regresní analýza, diskriminační analýza a logistická regrese. V poslední době jsou využívány rovněž neuronové sítě.

2.4.1 Regresní analýza

Regresní analýza je statistický nástroj, pomocí něhož je zkoumán vztah mezi nezávislou a závislou veličinou. Pokud je nezávislých proměnných více, hovoří se o vícenásobné regresní analýze, a základní vztah lze zapsat podle Hair (2014) jako

$$Y = X_1 + X_2 + \ldots + X_n, \tag{2.27}$$

kde Y je nezávislá, neboli vysvětlovaná proměnná a X_1 až X_n jsou jednotlivé vysvětlující proměnné. Jak nezávislá, tak závislá veličina jsou přitom spojité. Cílem regresní analýzy je využitím nezávislých faktorů odhadnout hodnotu závislé veličiny.

Model v regresní analýze může být lineární, kdy vztah mezi proměnnými je aditivní, navíc ještě lze modely dělit na linearitu v parametrech, v proměnných a v parametrech i proměnných. Druhou možností je nelineární model, který je možné převést využitím regresní mocninné a exponenciální funkce na lineární. Poslední možností je nelineární model, který nelze převést na lineární a není tak možné použít metodu nejmenších čtverců k odhadnutí koeficientů. Vícenásobný regresní model, který je lineární v parametrech i proměnných a je stochastický, což znamená, že zahrnuje náhodnou složku, je určen jako

$$Y = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + ... + \beta_n \cdot X_n + \varepsilon, \tag{2.28}$$

kde β_0 je odhadnuta úrovňová konstanta modelu, která určuje počátek regresní přímky, β_1 až β_n jsou odhadnuté parciální regresní koeficienty ε značí náhodnou složku. Parciální regresní koeficienty udávají, jak se změní hodnota vysvětlované proměnné, pokud se hodnota příslušné nezávislé proměnné změní o jednotku, za předpokladu, že hodnota ostatních vysvětlujících faktorů se nezmění. Význam náhodné složky je v zachycení náhodných vlivů, které nejsou v modelu obsaženy prostřednictvím dílčích proměnných.

Koeficienty jsou odhadovány pomocí metody nejmenších čtverců. Aby mohla být tato metoda použita, musí být podle Hančlová (2012) splněny tyto předpoklady:

- vysvětlující proměnné jsou nestochastické,
- střední hodnota náhodné složky je nulová,
- rozptyl náhodné složky je konečný a konstantní,
- náhodné složky jsou vzájemně nezávislé,
- kovariance mezi náhodnou složkou u a každou nezávislou proměnnou je nulová,
- náhodná složka má normální rozdělení,
- vysvětlující proměnné nejsou kolineární a
- model je správně specifikován.

Pokud tyto předpoklady nejsou splněny, pak je nutné proměnné, případně model, upravovat až do té doby, dokud nebudou splněny všechny předpoklady. Podrobnější popis této problematiky je uveden například v Hančlová (2012).

2.4.2 Diskriminační analýza

Vícenásobná regrese je nejčastěji používaná metoda pro odhad závislé proměnné, pokud závislá i nezávislá proměnná jsou spojité. V některých případech ale může být závislá hodnota kategorická a za těchto okolností je vhodnější použít diskriminační analýzu. Hlavním cílem diskriminační analýzy je správné přiřazení subjektů do jednotlivých tříd. V ekonomii a financích je typickým příkladem zařazení podniku buď do skupiny bonitních, nebo bankrotních podniků a to sestavením diskriminační analýzy s využitím poměrových ukazatelů společností. Na počátku analýzy jsou vybrány subjekty a rozděleny do tříd na bankrotující a bonitní a pro každou společnost jsou známé hodnoty nezávislých proměnných. S využitím těchto informací je sestavena výsledná diskriminační funkce, která je podle Hair (2014) dána jako

$$Z_j = a + W_1 \cdot X_{1j} + W_2 \cdot X_{2j} + ... + W_n \cdot X_{nj}, \qquad (2.29)$$

kde Z_{jk} je diskriminační *Z* skóre diskriminační funkce pro společnost *j*, a značí úrovňovou konstantu, W_n jsou diskriminační koeficienty a X_{nj} představují hodnoty nezávislých proměnných pro společnost *j*.

Výsledná hodnota *Z* skóre pro daný subjekt může nabývat hodnot (-∞; +∞) a je porovnávána s mezní hodnotou. Daný subjekt je podle tohoto principu zařazen mezi ty subjekty, jejichž *Z* skóre je podobné jako skóre daného subjektu. To znamená, že i hodnoty jednotlivých nezávislých proměnných jsou si podobné. Sestavený model obsahuje také určitou chybovost, z toho důvodu je možné kolem mezní hodnoty vymezit šedou zónu, a pokud *Z* skóre pro daný subjekt bude v tomto intervalu, tak není možné určit, do jaké třídy patří.

Aby diskriminační analýza mohla být správně provedena a odhadnutá diskriminační funkce byla vhodná, musí být podle Vochozka (2011) splněny tři omezující předpoklady:

- nezávislé proměnné odpovídají vícerozměrnému rozdělení pravděpodobnosti,
- kovarianční matice a matice závislostí mezi jednotlivými třídami subjektů jsou totožné a
- náklady chybné klasifikace jsou známé, to znamená, že když je určován mezní bod, pak by měla být zvažována pravděpodobnost chybného hodnocení pomocí diskriminační funkce.

2.4.3 Logistická regrese

Další využívanou statistickou metodou je logistická regrese, jejímž typickým znakem je kategoriální závislá proměnná a spojité nezávislé proměnné. Nejčastěji se přitom uvažuje s binární závislou proměnnou, která s pravděpodobností π nabývá hodnoty 1, pokud daná událost nastane a s pravděpodobností 1- π nabývá hodnoty 0, pokud daná událost nenastane. Pravděpodobnost výskytu události je v intervalu <0; 1>. Cílem logistické regrese je určit takové proměnné, které významně ovlivňují nezávislou proměnnou a odhadnutí logistického modelu, pomocí kterého je možné subjekty zařadit do jednotlivých skupin. Pro odhad logistických koeficientů je používána hodnota šance (odds) nebo logaritmus této šance.

Hodnota šance porovnává pravděpodobnost, že dojde k dané události s pravděpodobností, že k této události nedojde a je dána jako

$$\text{odds} = \frac{\pi}{1-\pi}. \tag{2.30}$$

Výsledná hodnota je z intervalu (0, ∞), abychom získali hodnotu z celého oboru reálných čísel, tak je potřeba hodnotu šance zlogaritmovat. Po zlogaritmování je možné získat výsledný logistický model, který lze zapsat jako

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 \cdot x_1 + ... + \beta_n \cdot x_n)}{1 + \exp(\beta_0 + \beta_1 \cdot x_1 + ... + \beta_n \cdot x_n)}. \tag{2.31}$$

Tento vztah lze převést do přijatelnější podoby logitovou transformací, na tzv. logit, $g(x)$, jehož hodnota je z intervalu (-∞; +∞) a lze jej podle Hair (2014) zapsat jako

$$g(x) = \ln(\frac{\pi}{1-\pi}) = \beta_0 + \beta_1 \cdot x_1 + ... + \beta_n \cdot x_n. \tag{2.32}$$

Na rozdíl od lineární regresní analýzy, kde je pro odhad koeficientů využívána metoda nejmenších čtverců, je v logistické regresi využívána pro odhad koeficientů metoda maximální věrohodnosti, a to kvůli nelineárnímu tvaru logistického modelu. Věrohodnostní funkce u logistické regrese, v případě, že závislá proměnná je dichotomická a tedy předpokladu binomického rozdělení y, je dána podle Hosmer a Lemeshow (2000) jako

$$L(\beta) = \sum_{i=1}^{n} \pi(x_i)^{y_i} \cdot (1 - \pi(x_i))^{1-y_i} \tag{2.33}$$

Tuto funkci je pro odhadnutí koeficientů nutné maximalizovat, přičemž z výpočetních důvodů je používán logaritmus této funkce, tedy

$$l(\beta) = \sum_{i=1}^{n} y_i \cdot \pi(x_i) + (1 - y_i) \cdot (1 - \pi(x_i)) \tag{2.34}$$

Po derivaci této funkce jsou zjištěny hodnoty odhadnutých koeficientů, které udávají, jak daná vysvětlující proměnná ovlivňuje nezávislou proměnnou.

Nevýhodou logistické regrese a diskriminační analýzy je mimo jiné to, že musí být známy hodnoty závislých proměnných pro každou společnost. To znamená, že pokud by bylo potřeba zjistit vliv nezávislých proměnných na dobu do výskytu bankrotu u podniku, musel by být známý časový okamžik, kdy k bankrotu došlo a přesný okamžik, kdy k úpadku nedošlo. V praxi však je k dispozici údaj pouze o době, kdy nastal bankrot, avšak pokud k úpadku společnosti nedošlo a podnik dále funguje, tak není možné určit hodnotu závislé proměnné. Tento problém je odstraněn pomocí cenzorování v analýze přežití, která je detailně popsána v následující kapitole.

3 Teoretické vymezení analýzy přežití

Obsahem této kapitoly je popsání analýzy přežití, což je celý soubor statistických metod, které se snaží určitým způsobem zachytit dobu mezi dvěma událostmi. Vstupními údaji nutnými pro analýzu přežití jsou počáteční a koncová událost a vysvětlující proměnné, které mají vliv na to, zda a kdy koncová událost nastane. Nejčastěji je tato analýza využívána v oblasti medicíny, kde je sledován čas od výskytu nemoci do úmrtí pacienta. Další možné využití je ve strojírenství, pro demografické nebo sociální výzkumy a v ekonomii, kde bývá modelována například doba od ztráty zaměstnání do nalezení nového zaměstnání. V této práci je aplikováno využití ve financích, a to zkoumáním doby do bankrotu daného subjektu. To je vhodné především pro banky a další instituce, které zapůjčují peníze a mohou tak namodelovat, s jakou pravděpodobností v době splatnosti půjčky nebude subjekt v bankrotu.

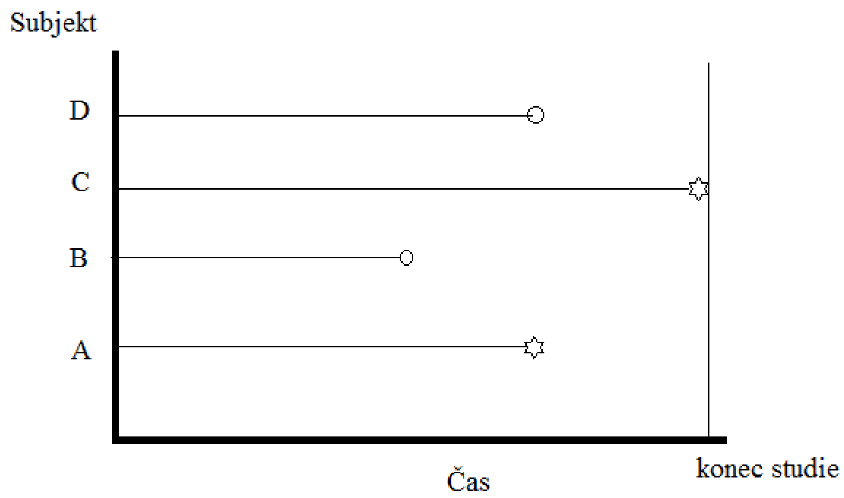
První část kapitoly je věnována základním pojmům v analýze přežití, jako je cenzorování, funkce přežití nebo hazardní funkce. V další části jsou popsány používané metody, které jsou členěny na neparametrické, semiparametrické a parametrické. Na závěr kapitoly je vysvětlen princip sestavení modelu a jeho verifikace, a to jak statistická významnost odhadnutých parametrů, tak celková adekvátnost modelu.

3.1 Cenzorování dat

Výhodou analýzy přežití na rozdíl od běžných statistických nástrojů je schopnost pracovat s neúplnými daty. Obvykle totiž není z různých důvodů možné sledovat všechny subjekty ve studii až do koncové události. Rozlišováno je podle Hosmer (2008) cenzorování zleva, zprava a intervalové cenzorování. K cenzorování zleva dochází v případě, že u subjektu došlo ke sledované události ještě před začátkem studie. Intervalové cenzorování nastává, pokud nelze přesně určit dobu přežití, jelikož nebylo možné subjekt neustále sledovat a k události došlo někdy mezi dvěma pozorováními. Nejčastěji však dochází k cenzorování zprava, kdy není známa skutečná doba přežití a to buď kvůli tomu, že ke koncové události během sledovaného období nedošlo, nebo z důvodu ztráty informací o daném subjektu během sledovacího období.

Základními údaji, které jsou k dispozici pro každý sledovaný subjekt, je tedy čas T a časový cenzor c. Čas T v případě necenzorovaných dat udává dobu od počátku sledování do výskytu koncové události. U cenzorovaných subjektů pak T udává dobu od počátku do ukončení pozorování z důvodu ztráty informací, nebo konce studie, kdy u subjektu stále nedošlo ke koncové události. Pro rozlišení cenzorovaných a necenzorovaných dat je využíváno značení $c = 0$ pro cenzorovaná pozorovaní a $c = 1$ pro necenzorovaná pozorování. Na obrázku (Obrázek 3.1) je příklad cenzorování zprava u subjektu A, kde došlo k odebrání ze studie v důsledku nedostatku informací, a u subjektu C, kde nedošlo ke koncové události do ukončení studie. Subjekty B a D jsou necenzorované, je znám přesný časový okamžik, kdy došlo ke sledované události.

Obrázek 3.1 Příklad cenzorování zprava



Zdroj: vlastní vypracování

3.2 Funkce přežití a hazardní funkce

Základními funkcemi, které v analýze přežití umožňují prvotní náhled na získaná data, je funkce přežití a hazardní funkce. Funkce přežití je inverzní funkcí k distribuční funkci a udává pravděpodobnost, že subjekt přežije alespoň do času t, přičemž ji je možné definovat podle Cleves (2010) jako

$$S(t) = 1 - F(t) = \Pr(T > t), \tag{3.1}$$

kde $S(t)$ je funkce přežití, $F(t)$ je distribuční funkce, přičemž hodnota funkce v čase t vyjadřuje pravděpodobnost, že doba přežití nebude větší než čas t a T představuje čas do výskytu koncové události.

Druhou základní funkci je hazardní funkce, kterou je možné definovat podle Hosmer (2008) jako pravděpodobnost, že dojde ke koncové události krátce po čase t za předpokladu, že subjekt dožil do tohoto časového okamžiku. Vztah pro výpočet této funkce je možné zapsat následujícím způsobem

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t + \Delta t > T > t | T > t)}{\Delta t}, \tag{3.2}$$

kde $h(t)$ udává hazardní funkci.

Dále je možné dokázat vztah mezi hazardní funkcí, funkcí přežití a distribuční funkcí, který lze zapsat jako

$$h(t) = \frac{F(t)}{S(t)}. \tag{3.3}$$

Z hazardní funkce lze rovněž odvodit kumulativní hazardní funkci, která vyjadřuje pravděpodobnost výskytu události od počátku sledování do času t a je tak možné ji definovat pomocí hazardní funkce jako

$$H(t) = \int_0^t h(u)du, \tag{3.4}$$

přičemž $H(t)$ značí kumulativní hazardní funkci, která kromě pravděpodobnosti výskytu koncové události od počátku do času t udává také počet výskytů koncových události za daný časový interval. Tuto funkci je možné odvodit také z funkce přežití, kterou lze s využitím kumulativní hazardní funkce zapsat jako

$$S(t) = \exp(-H(t)),\ t \geq 0. \tag{3.5}$$

Pomocí těchto vztahů je tedy možné odvodit funkci přežití, hazardní funkci a distribuční funkci, pokud je známa alespoň jedna z těchto funkcí. Pro tyto účely je nutné znát rozdělení pravděpodobnosti náhodné veličiny, což je v tomto případě čas T. Podle toho, jaké předpoklady o rozdělení pravděpodobnosti je nutné určit, jsou v analýze přežití rozlišovány tři základní skupiny metod, a to:

- neparametrické metody, u kterých není nutné určit žádný předpoklad rozdělení pravděpodobnost doby přežití a nelze zkoumat vliv vysvětlujících proměnných,
- semiparametrické metody, u kterých také není nutné definovat předpoklad o rozdělení pravděpodobnosti, ale je možné zahrnout vysvětlující proměnné, jejichž odhady mají parametrickou strukturu a
- parametrické metody, pro jejichž aplikaci je potřebné určit rozdělení pravděpodobnosti doby přežití, přičemž mezi využívané rozdělení pravděpodobnosti patří exponenciální, Weibullovo, log-normální nebo zobecněné gamma rozdělení.

3.3 Neparametrické metody

První skupinou metod v analýze přežití jsou neparametrické metody, které nevyžadují žádný předpoklad o rozdělení pravděpodobnosti a nezahrnují vysvětlující proměnné. Analyzují tedy pouze čas mezi počáteční a koncovou událostí a to využitím Kaplan–Meierova odhadu funkce přežití nebo Nelson–Aalenova odhadu. Využití těchto metod spočívá v grafickém zobrazení jednotlivých funkcí, a porovnání jednotlivých funkcí pro více skupin. Toto porovnání mezi více skupinami lze provést nejen graficky, ale i analyticky využitím různých testů, například log – rank testu nebo obecného Wilcoxonova testu.

Kaplan–Meierův odhad, jak uvádí Hosmer (2008), slouží k odhadu funkce přežití v každém daném okamžiku, ve kterém došlo ke koncové události, což je například bankrot společnosti. Metoda zahrnuje jak cenzorovaná, tak necenzorovaná data a funguje na principu podmíněné pravděpodobnosti. K odhadu funkce přežití je zapotřebí znát celkový počet subjektů v riziku v čase t_i (n_i) a počet subjektů, u kterých došlo ke koncové události v čase t_i (d_i), je možné jej zapsat ve tvaru

$$\hat{S}_t = \prod_{t_i \le t} \frac{n_i - d_i}{n_i}. \tag{3.6}$$

Výsledkem je tzv.“schodovitá“ funkce, která je konstantní mezi dvěma časy úpadku, přičemž jednotlivé schody určují čas, kdy došlo ke koncové události. Cenzorovaná data ovlivňují výši schodu, neboť určují počet subjektů, které podléhají riziku v daném časovém okamžiku. Z výše uvedeného je možné odvodit, že $\hat{S}_t = 1$ až do prvního výskytu sledované události a dále funkce klesá, maximálně však do nuly.

Druhou možností, jak odhadnout funkci přežití, je Nelson–Aalenův odhad, který je založen na odhadu kumulativní hazardní funkce a z něj je poté odhadnuta funkce přežití. Odhad kumulativní funkce je dán jako

$$\hat{H}(t) = \sum_{t_i \le t} \frac{d_i}{n_i}. \tag{3.7}$$

Po odhadnutí kumulativní hazardní funkce lze dosazením do vzorce (3.5) vypočítat odhad funkce přežití způsobem

$$\hat{S}(t) = \exp(-\hat{H}(t)). \tag{3.8}$$

Rozdíl mezi Kaplan–Meierovým a Nelson–Aalenovým odhadem funkce přežití je v počítání s cenzorovanými daty. Zatímco u prvního zmíněného odhadu jsou do rizikové skupiny zahrnuty cenzorované časy až do okamžiku cenzorování včetně, u Nelson–Aalenova odhadu kumulativní hazardní funkce jsou cenzorované časy v daném okamžiku vyloučeny. Z toho vyplývá, že Kaplan–Meierův odhad je vždy menší nebo roven Nelson–Aalenovu odhadu funkce přežití. Při dostatečně velkém vzorku dat však dle Gill (1980) mezi těmito odhady není významný rozdíl.

Neparametrické metody jsou užitečné při porovnání doby přežití mezi dvěma nebo více skupinami subjektů. Základní rozdíly lze vyčíst z grafického znázornění funkce přežití nebo kumulativní hazardní funkce. Pokud se tyto funkce liší, pak je nutné určit, zda je rozdíl mezi jednotlivými skupinami statisticky významný. Pro tyto účely slouží různé statistické testy, jejichž výpočet je založen na kontingenčních tabulkách a porovnávání skutečného a očekávaného počtu výskytu koncových událostí. Nulová hypotéza testů je ve tvaru: „funkce přežití jsou v porovnávaných skupinách shodné". Při srovnávání dvou skupin je možné označit jednu skupinu jako 0 a druhou jako 1, očekávaný počet úpadků ve skupině 1 (e_{1i}) pak je definován jako

$$\hat{e}_{1i} = \frac{n_{1i} \cdot d_i}{n_i}, \tag{3.9}$$

kde n_{1i} je počet subjektů v riziku ve skupině 1, n_i značí celkový počet subjektů v riziku a d_i představuje počet výskytů koncových událostí ve skupině 1 (d_{1i}) a ve skupině 0 (d_{0i}), tedy

$$d_i = d_{1i} + d_{0i}. \tag{3.10}$$

Výpočet testové statistiky je dle Hosmer (2008) dán vztahem

$$Q = \frac{\left[\sum_{i=1}^{m} w_i (d_{1i} - \hat{e}_{1i})\right]^2}{\sum_{i=1}^{m} w_i^2 \cdot \hat{v}_{1i}}. \tag{3.11}$$

přičemž m je celkový počet výskytu koncových událostí, w_i značí váhu, která se liší v závislosti na konkrétním použitém testu. Pro logrank test, který je použit v této práci, je w_i rovno 1. Pro obecný Wilcoxonoův test je $w_i = n_i$ a rovněž další testy se liší pouze hodnotou této váhy. Poslední neznámou v rovnici (3.10) je $\hat{v}_{1i}$, což je odhad rozptylu d_{1i} určen na základě hypergeometrického rozdělení jako

$$\hat{v}_{1i} = \frac{n_{1i} \cdot n_{0i} \cdot d_i \cdot (n_i - d_i)}{n_i^2 \cdot (n_i - 1)}. \tag{3.12}$$

Vypočítána testová statistika (3.11) odpovídá za předpokladu platnosti nulové hypotézy, že funkce přežití jsou stejné, a dostatečně velkého vzorku pozorování i koncových událostí χ^2 rozdělení pravděpodobnosti s jedním stupněm volnosti.

3.4 Coxův proporciální hazardní model

Neparametrické metody jsou užitečné při porovnávání několika skupin, ovšem jejich nevýhodou je, že do nich nelze zahrnout vliv vysvětlujících proměnných. Semiparametrické modely, z nichž nejvyužívanější je Coxův model, jsou charakteristické zahrnutím nezávislých proměnných a přitom není potřeba určit předpoklad o rozdělení pravděpodobnosti doby přežití. Výhodou tohoto zjednodušení je poměrně rychlé zjištění vlivu jednotlivých faktorů na funkci přežití a hazardní funkci a také určení hazardního poměru.

3.4.1 Funkce přežití a hazardní funkce v Coxově modelu

Regresní model, který navrhl Cox (1972) je ve tvaru

$$h(t,x,\beta) = h_0(t) \cdot \exp(x\beta_x), \tag{3.13}$$

kde $h(t,x,\beta)$ je hazardní funkce, která je dána jakou součin $h_0(t)$, což je základní hazardní funkce, která je závislá pouze na čase, a $\exp(x\beta_x)$, kde β_x jsou odhadované koeficienty a x je vysvětlující faktor. Tato část funkce není závislá na čase, ale udává, jak ovlivňují nezávislé proměnné hazardní funkci.

Pro lepší interpretaci výsledků je místo koeficientů využíván hazardní poměr (*HR*), který je v Cleves (2010) popisován jako změna rizika, že nastane koncová událost při změně dané proměnné o jednotku a je definován jako

$$HR = \exp(\beta). \tag{3.14}$$

Přičemž pokud je hazardní poměr menší než 1, tak nárůst hodnoty ukazatele vede ke snížení rizika výskytu koncové události. V opačném případě, tedy když hodnota hazardního poměru je vyšší než 1, tak čím větší hodnota vysvětlující proměnné, tím větší riziko výskytu koncové události.

Pokud je známa hazardní funkce Coxova modelu, pak dosazením do rovnice (3.8) lze získat funkci přežití se zahrnutím vysvětlujících proměnných, a to pomocí vztahu

$$S(t,x,\beta) = \exp(-H(t,x,\beta)), \tag{3.15}$$

kde $H(t,x,\beta)$ značí kumulativní hazardní funkci s nezávislou proměnnou x a odhadem parametru této proměnné. Vzhledem k tomu, že kumulativní hazardní funkce je součtem hazardních funkcí, tak ji v tomto případě lze vyjádřit s využitím vztahu (3.4) jako

$$H(t,x,\beta) = \int_0^t h(u,x,\beta)du = \int_0^t h_0(u)\exp(x\beta)du = H_0(\mathrm{t})\exp(x\beta). \tag{3.16}$$

Kumulativní hazardní funkce je obdobně jako hazardní funkce dána součinem základní kumulativní hazardní funkce, která je závislá na čase a udává míru rizika, která je shodná pro všechny subjekty. Druhým faktorem je $\exp(x\beta)$, který upravuje kumulativní hazardní funkci podle vysvětlujících proměnných. Dosazením rovnice (3.16) do rovnice (3.15) je získána funkce přežití u Coxova hazardního modelu ve tvaru

$$\mathrm{S}(t,x,\beta)=\exp(-H_0(\mathrm{t})\exp(x\beta))=\left[S_0(\mathrm{t})\right]^{\exp(x\beta)}. \tag{3.17}$$

Přičemž $S_0(\mathrm{t})=\exp(-H_0(\mathrm{t}))$ a je to základní funkce přežití, která nabývá oboru hodnot <0,1>.

3.4.2 Odhad parametrů modelu

Běžně používanou metodou pro odhad parametrů modelu je metoda maximální věrohodnosti. Nutnou podmínkou pro aplikaci této metody je schopnost určit rozdělení pravděpodobnosti náhodné veličiny, což u Coxova hazardního modelu není splněno. Byla proto Coxem (1975) navrhnuta metoda parciální maximální věrohodnosti, přičemž věrohodnostní funkce je závislá jen na hledaných parametrech a není závislá na základní hazardní funkci, tedy náhodné veličině. Předpokladem je, že odhadnuté parametry z této funkce mají stejné rozdělení pravděpodobnosti, jako odhady z plně maximalizované věrohodnostní funkce.

Parciální věrohodnostní funkce jsou rozdílné v závislosti na tom, zda došlo v daném časovém okamžiku k výskytu konečné události u jednoho subjektu, nebo u více. V případě většího počtu výskytu sledovaných událostí v jednom časovém okamžiku, je nutné parciální věrohodnostní funkci odhadnout, například Breslowovou nebo Efronovou aproximací, jak uvádí Hosmer (2008). Pokud k výskytu sledované události došlo v daném čase právě u jednoho subjektu, pak je parciální věrohodnostní funkce $L_p(\beta)$ definována dle Hosmer (2008) jako

$$L_p(\beta)=\prod_{i=1}^{m}\frac{\exp(x_i\beta)}{\sum\limits_{j\in R(t_{(i)})}\exp(x_i\beta)}, \tag{3.18}$$

kde $R(t_{(i)})$ značí množinu všech subjektů v riziku v čase $(t_{(i)})$ a m jsou veškeré pozorované časy výskytu koncových událostí seřazených od nejmenšího. Po zlogaritmování funkce (3.18) lze získat log – parciální věrohodnostní funkci $l_p(\beta)$ ve tvaru

$$l_p(\beta) = \sum_{i=1}^{m} \left\{ x_i\beta - \ln\left[\sum_{j \in R(t_{(i)})} \exp(x_i\beta) \right] \right\}. \tag{3.19}$$

Pro získání odhadů parametrů je nutné nalézt maximum log – parciální věrohodnostní funkce, k čemuž slouží derivace pravé strany rovnice (3.19) podle parametrů β, což lze vyjádřit jako

$$\frac{\partial l_p(\beta)}{\partial \beta} = \sum_{i=1}^{m} \left\{ x_i - \frac{\sum_{j \in R(t_{(i)})} x_j \exp(x_j\beta)}{\sum_{j \in R(t_{(i)})} \exp(x_j\beta)} \right\}, \tag{3.20}$$

dále proložení nulou a vyřešením soustavy k rovnic o k neznámých, přičemž m značí počet hledaných parametrů. K tomuto účelu slouží více algoritmů, v aplikační části práce je využita Newton–Raphsonova iterační metoda, která je popsána více například v Collet (2003).

Rozptyl odhadnutých parametrů je vypočítán z hodnoty druhé derivace log – parciální věrohodnostní funkce podle hledaných parametrů β, kterou po úpravách lze zapsat jako

$$\frac{\partial^2 l_p(\beta)}{\partial \beta^2} = -\sum_{i=1}^{m} \sum_{j \in R(t_{(i)})} w_{ij}(\mathrm{x}_j - \bar{x}_{wi})^2, \tag{3.21}$$

kde

$$w_{ij} = \frac{\exp(x_j\beta)}{\sum_{i \in R(t_{(i)})} \exp(x_i\beta)}, \tag{3.22}$$

a

$$\bar{x}_{wi} = \sum_{j \in R(t_{(i)})} w_{ij} \cdot x_j. \tag{3.23}$$

Ke zjištění rozptylu odhadnutých parametrů je zapotřebí záporná hodnota druhé derivace, která bývá označována jako pozorovaná informační matice (*I(β)*) a je tedy rovna

$$I(\beta) = -\frac{\partial^2 l_p(\beta)}{\partial \beta^2}. \tag{3.24}$$

Odhad kovarianční matice je roven inverzní hodnotě pozorované informační matice a odhady rozptylu jednotlivých koeficientů leží na diagonále této matice, kterou lze zapsat jako

$$Var(\hat{\beta}) = I(\hat{\beta})^{-1}. \tag{3.25}$$

Směrodatná odchylka odhadu parametrů $s(\hat{\beta})$ je dána jako odmocnina z rozptylu, to znamená

$$s(\hat{\beta}) = \sqrt{Var(\hat{\beta})}. \tag{3.26}$$

Ze směrodatné odchylky lze odvodit standardní neboli směrodatnou chybu odhadu parametru, která kvantifikuje nepřesnost měření a závisí také na počtu prvků v souboru n a je možné ji vyjádřit jako

$$SE(\hat{\beta}) = \frac{s(\hat{\beta})}{\sqrt{n}}. \tag{3.27}$$

To znamená, že s rostoucím souborem klesá směrodatná chyba odhadu parametru. Směrodatnou chybu je možné mimo jiné využít pro určení intervalu spolehlivosti, kde koncové body těchto intervalů jsou určeny jako

$$\hat{\beta} \pm z_{1-\alpha/2} \cdot SE(\hat{\beta}), \tag{3.28}$$

kde α je hladina významnosti, pro kterou je interval spolehlivosti hledán a $z_{1-\alpha/2}$ je hodnota inverzní funkce k distribuční funkci normovaného normálního rozdělení pro danou hladinu významnosti.

3.5 Parametrické metody

Poslední skupinou metod využívanou v analýze přežití jsou parametrické metody. Výhodou těchto modelů je přesnější odhad parametrů modelu a umožnění predikce času přežití jednotlivých subjektů, pokud lze odhadnout správně tvar základní hazardní funkce. S tím souvisí odhad parametrů, který je prováděn využitím metody plné maximální věrohodnosti, namísto metody parciální maximální věrohodnosti, která je používána při odhadu v Coxově modelu. Nejčastějšími rozděleními, kterými se řídí doba přežití, jsou exponenciální, Weibullovo, log-logistické nebo zobecněné Gamma rozdělení. Dále je u parametrických modelů rozlišováno mezi proporciálním hazardním (PH) modelem, což znamená, že hazardní poměr je konstantní v čase a je závislý pouze na hodnotě vysvětlujících proměnných a modelem zrychleného času (AFT), kde jsou vysvětlující proměné vztaženy k rizikové funkci náhodné veličiny. V této práci je dále pracováno s PH modelem, z důvodu porovnání se semiparametrickým modelem, kde hazardní poměr také není závislý na čase. Parametrický PH model je možné sestavit pouze pro různá rozdělení pravděpodobnosti, jedním z často používaných je rozdělení Weibullovo, které je použito také v tomto případě.

3.5.1 Funkce přežití a hazardní funkce v parametrickém modelu

Pokud se tedy základní hazardní funkce řídí Weibullovým rozdělením, pak je definována podle Cleves (2010) jako

$$h_0 = \lambda \cdot p \cdot t^{p-1}, \tag{3.29}$$

kde λ, $p > 0$, přičemž λ je parametr měřítka a p je parametr tvaru. Dále platí, že pokud je $p < 1$, pak je hazardní funkce klesající, pokud je $p > 1$, pak je tato funkce rostoucí a při $p = 1$ je model možné zredukovat na exponenciální a hazardní funkce je konstantní. Parametr měřítka λ je možné vyjádřit jako

$$\lambda = \exp(\beta_0), \tag{3.30}$$

to znamená exponent odhadnuté konstanty modelu.

Po dosazení základní hazardní funkce, která odpovídá Weibullovému rozdělení (3.29) do rovnice (3.13) vyjde vztah pro hazardní funkci za předpokladu Weibullova rozdělení následovně

$$h(t, x, \beta, \lambda, p) = \lambda \cdot p \cdot t^{p-1} \cdot \exp(x\beta_x) = p \cdot t^{p-1} \cdot \exp(\beta_0 + x\beta_x). \tag{3.31}$$

Z této hazardní funkce je zřejmé, že parametr tvaru p je pro všechny subjekty stejný a tedy i hazardní funkce je stejná. Oproti tomu parametr měřítka λ je ovlivněn jednotlivými faktory, které jsou rozdílné pro dané subjekty, a tak i hazardní funkce je rozdílná pro různé subjekty.

Při platnosti rovnice (3.15) je určena funkce přežití pro parametrický model, kde čas odpovídá Weibullovu rozdělení, pomocí (3.31) jako

$$S(t, x, \beta, \lambda, p) = \exp(-\exp(x\beta_x) \cdot \lambda \cdot t^p). \tag{3.32}$$

3.5.2 Odhad koeficientů parametrického modelu

Pokud je určen předpoklad, kterému rozdělení pravděpodobnosti odpovídá základní hazardní funkce, tak už je možné použit pro odhad hledaných parametrů metodu maximální věrohodnosti. Tato metoda je přesnější než používaná metoda parciální věrohodnosti v Coxově modelu. Problémem však v analýze přežití je výskyt cenzorovaných dat. To se týká těch subjektů, u kterých v daném intervalu, ve kterém pobíhala studie, nedošlo k výskytu koncové události. Není tedy znám přesný čas, po který subjekt přežil, ovšem je známo do jaké doby alespoň nedošlo k výskytu sledované události. Do věrohodnostní funkce je tak zahrnuta pravděpodobnost, že u firmy s vektorem vysvětlujících proměnných x nedojde ke koncové události alespoň do času t, což je vystihnuto funkcí přežití. Pokud je známé datum bankrotu, pak se jedná o necenzorovaná data a do věrohodnostní funkce vstupuje pravděpodobnost, že společnost s daným vektorem nezávislých proměnných x zbankrotuje v čase t, vyjádřena pomocí hustoty pravděpodobnosti $f(t, x_i, \beta)$.

Hledanými parametry jsou odhadované koeficienty β, dále parametr měřítka λ a parametr tvaru p. Při definování časového cenzoru $c_i = 1$ pro necenzorovaná data a $c_i = 0$ pro cenzorovaná data, je věrohodnostní funkce vyjádřena podle Hosmer (2008) tímto způsobem

$$L = \prod_{i=1}^{m} \left\{ \left[f(t, x_i, \beta) \right]^{c_i} \cdot \left[S(t, x_i, \beta) \right]^{1-c_i} \right\}. \tag{3.33}$$

Pro maximalizaci je využíván logaritmický výraz rovnice, tedy

$$l = \sum_{i=1}^{m} \left\{ c_i \cdot \ln\left[f(t, x_i, \beta) \right] + (1 - c_i) \cdot \ln\left[S(t, x_i, \beta) \right] \right\}. \tag{3.34}$$

Když z rovnice (3.3) je vyjádřena funkce hustoty jako součin hazardní funkce a funkce přežití a na tomto základě upravena rovnice (3.34), pak je získán výraz

$$l = \sum_{i=1}^{m} \left\{ c_i \cdot \ln\left[h(t, x_i, \beta) \right] + \ln\left[S(t, x_i, \beta) \right] \right\}. \tag{3.35}$$

Dosazením hazardní funkce (3.31) a funkce přežití parametrického modelu (3.32), do rovnice (3.35) a úpravě, je získán logaritmus věrohodnostní funkce pro dané rozložení, tedy

$$l = \sum_{i=1}^{m} \left\{ c_i \cdot \ln\left[\ln(\lambda \cdot p) + (p-1) \cdot \ln(t) + (x\beta_x) \right] - \exp(x\beta_x) \cdot \lambda \cdot t^p \right\}. \tag{3.36}$$

Tuto rovnici lze podle Collet (2003) zjednodušit o vynechání výrazu $\sum_{i=1}^{n} c_i \cdot \ln(t)$, jelikož neobsahuje žádný hledaný parametr. Pro získání odhadovaných parametrů tak je nutné maximalizovat logaritmus věrohodnostní funkce

$$l = \sum_{i=1}^{m} \left\{ c_i \cdot \ln\left[\ln(\lambda \cdot p) + p \cdot \ln(t) + (x\beta_x) \right] - \exp(x\beta_x) \cdot \lambda \cdot t^p \right\}, \tag{3.37}$$

a to vzhledem k odhadovaným parametrům, tedy koeficientům β_x, dále parametru λ a parametru p. Tento proces bude podobně jako u Coxova modelu proveden využitím Newton-Raphsonovy iterační metody.

Kovarianční matici odhadnutých koeficientů je možné získat z inverzní matice k Fisherově informační matici. Tato matice je definována jako druhá derivace log-věrohodnostní funkce podle hledaných parametrů, kovarianční matici tedy lze zapsat jako

$$\begin{pmatrix} Var(\beta) & Cov(\beta, p) \\ Cov(\beta, p) & Var(p) \end{pmatrix} = \begin{pmatrix} \dfrac{\partial^2 l}{\partial \beta^2} & \dfrac{\partial^2 l}{\partial \beta \partial p} \\ \dfrac{\partial^2 l}{\partial \beta \partial p} & \dfrac{\partial^2 l}{\partial p^2} \end{pmatrix}^{-1}. \tag{3.38}$$

Pokud je znám rozptyl odhadnutých parametrů, tak vzhledem k tomu, že koeficienty byly odhadnuty metodou maximální věrohodnosti, jsou tyto odhady asymptoticky normální a je možné získat intervaly spolehlivosti, které jsou podle Nelson (1982) definované jako

$$\theta_h = \theta \cdot \exp(\frac{K_\alpha \cdot \sqrt{\theta}}{\theta}), \tag{3.39}$$

pro horní interval a

$$\theta_d = \frac{\theta}{\exp(\frac{K_\alpha \cdot \sqrt{\theta}}{\theta})}, \tag{3.40}$$

pro spodní interval na hladině významnosti α. Intervaly spolehlivosti mohou být hledány pro odhadnuté koeficienty β a pro parametr *p*, znak θ je tak roven buď β, nebo parametru *p*, podle toho pro co je interval spolehlivost počítán. Poslední neznámou je K_α, která může být určena využitím rovnice

$$\alpha = \frac{1}{\sqrt{2\pi}} \int_{K_\alpha}^{\infty} \exp(\frac{-t^2}{2}) dt = 1 - \phi(K_\alpha), \tag{3.41}$$

kde ϕ je distribuční funkce normovaného normálního rozdělení.

3.6 Testování významnosti odhadnutých koeficientů

Po odhadnutí regresních koeficientů je vhodné zjistit, zda jsou tyto koeficienty statisticky významné. K tomuto účelu je možné dle Hosmer (2008) využít tří různých přístupů, konkrétně se jedná o test poměru věrohodnostní funkce, Waldovu statistiku a skórový test.

První z uvedených testů je založen na dvojnásobku rozdílu mezi hodnotou log-věrohodnostní funkce, případně log-parciální věrohodnostní funkce s příslušnými nezávislými veličinami a hodnotou této funkce bez nezávislých veličin. Testová statistika je dána jako

$$G = 2(L_p(\hat{\beta}) - L_p(0)), \tag{3.42}$$

kde *G* je testová statistika, která odpovídá χ^2 rozdělení pravděpodobnosti s *k* stupni volnosti, přičemž *k* udává počet odhadnutých parametrů. Hodnota log–parciální věrohodnostní funkce bez nezávislých veličin je značena jako $L_p(0)$ a je možné ji zapsat vztahem

$$L_p(0) = -\sum_{i=1}^{m} \ln(n_i), \tag{3.43}$$

kde n_i udává počet subjektů v rizikové skupině v čase $t_{(i)}$. Nulová hypotéza testu zní

$$H_0 : \hat{\beta}_1 = \hat{\beta}_2 = ... = \hat{\beta}_p = 0, \tag{3.44}$$

alternativní hypotéza je negací, tedy alespoň jeden z koeficientů je nenulový.

Druhou možností, jak zjistit významnost odhadnutých koeficientů je Waldova statistika, která vyjadřuje poměr mezi odhadnutým koeficientem a směrodatné chyby tohoto koeficientu, tedy

$$z = \frac{\hat{\beta}}{SE(\hat{\beta})}, \tag{3.45}$$

Při platnosti nulové hypotézy, že koeficient je nenulový, odpovídá testová statistika normálnímu rozdělení.

Posledním využívanou metodou pro zjištění významnosti je skórový test, kterým je také možné zkoumat významnost každého koeficientu zvlášť, přičemž testová statistika je rovna

$$z^* = \frac{\frac{\partial L_p(\beta)}{\partial \beta}}{\sqrt{I(\beta)}}, \tag{3.46}$$

tedy poměr mezi derivací log-parciální věrohodnostní funkce (3.21) a odmocninou pozorované informační matice (3.24). Testová statistika za předpokladu nulové hypotézy má opět normální rozdělení pravděpodobnosti.

Se statistickým testováním hypotéz souvisí také pojem p-hodnota. V současnosti počítačové programy často uvádějí právě tuto hodnotu místo testové statistiky a tak je vhodné definovat, co tato hodnota znamená. P-hodnotu je možné interpretovat jako nejmenší hladinu významnosti, na které lze zamítnout nulovou hypotézu. Pro posouzení, zda je možné přijmout nulovou hypotézu na dané hladině významnosti je potřeba porovnat zjištěnou p-hodnotu se zvolenou hladinou významnosti α. Přičemž mohou nastat dvě následující situace:

a) p-hodnota $\leq \alpha$, pak je nulová hypotéza zamítnuta a lze přijmout alternativní hypotézu, nebo
b) p-hodnota $> \alpha$, pak je nulová hypotéza přijata.

3.7 Sestavení modelu

Zařazení příslušných proměnných do výsledného modelu funguje na stejném principu jak u Coxova modelu, tak u parametrického modelu. Na počátku je k dispozici obvykle více potenciálních proměnných, než je nutné do modelu zařadit, neboť některé z nich jsou statisticky nevýznamné a pouze nepříznivě ovlivňují celkový model. Existuje mnoho metod, jak vybrat dané nezávislé veličiny, například metoda zpětné eliminace, dopředního výběru nebo jejich kombinace. Využitou metodou pro sestavení modelu v této práci je metoda navržena autorem Collet (2003), kdy je kombinována metoda dopředního výběru a zpětné eliminace a o každé zařazené proměnné rozhoduje sestavovatel, nikoli statistický software. Základem pro posouzení vysvětlujících faktorů je test poměru parciální, respektive plnohodnotné věrohodnostní funkce a Waldova statistika, obě metody jsou popsány v předcházející podkapitole (3.6). Tvorbu výsledného modelu je možné rozčlenit do několika fází.

Prvním krokem je vytvoření jednorozměrných modelů, kdy každá vysvětlující proměnná je do modelu zařazena samostatně. Do další fáze pak jsou zařazeny takové veličiny, které jsou statisticky významné alespoň na 20% hladině významnosti například pomocí Waldovy statistiky.

Dále je sestaven vícerozměrný model s proměnnými, které se v prvním kroku ukázaly jako významné a využita metoda zpětné eliminace pro odstranění nevýznamných proměnných. To znamená, že pokud jsou v tomto modelu některé veličiny statisticky nevýznamné na dané hladině významnosti, například 10%, pak je odebrána ta proměnná, která je nejméně významná podle Waldovy statistiky. Vhodné je taky otestovat původní model s modelem bez vyřazené proměnné pomocí testu poměru věrohodnostní funkce. Když i tento test potvrdí nevýznamnost proměnné, je možné ji odebrat z modelu. Tento postup je opakován do té doby, dokud je v modelu statisticky nevýznamná proměnná.

Některé proměnné, které se v prvotní analýze zdály jako nevýznamné, mohou být významné ve vytvořeném modelu. Dalším krokem tedy je zařazení postupně všech proměnných, které byly v prvním kroku vyloučeny a zjištění, zda nejsou významné v tomto modelu. V této fázi je nutné sledovat, zda proměnná, která by byla přidána do modelu, negativně neovlivňuje ty proměnné, které již v modelu jsou. Pokud by tedy po přidání proměnné došlo k výraznému snížení významnosti některé dříve zařazené proměnné, tak by nová veličina neměla být do modelu přidána. Po tomto kroku už jsou známy všechny veličiny, které mají významný vliv na nezávislou proměnnou. Tímto postupem vznikne výsledný model.

3.8 Hodnocení adekvátnosti modelu

Po úspěšném sestavení modelu je vždy potřeba ohodnotit, zda je model vypovídající a jestli jsou splněny podmínky pro využití odpovídající metody. V případě analýzy přežití hrají velkou roli cenzorovaná data, a proto není možné použít pouze grafické metody nebo metody postačující v lineárním regresním modelování. Běžně používanými metodami jsou postupy založené na reziduích, které je možné vypočítat pro každý pozorovaný subjekt. Výběr a využití reziduí se může lišit podle typu dat a preferencí sestavovatele modelu, obvykle se ovšem používají čtyři typy reziduí. Aplikací Schoenfeldových reziduí je možné zkontrolovat podmínku proporcionality hazardu. Cox-Snellova rezidua jsou využívána pro hodnocení celkové adekvátnosti modelu. Martingalova a deviantní rezidua jsou vhodné pro identifikování subjektů, které nebyly modelem příliš vhodně zachyceny nebo takových subjektů, které mají nepřiměřený vliv na hodnotu odhadnutých parametrů. Podrobnější popsání a odvození následujících vztahů lze nalézt v Collet (2003), odkud byly čerpány teoretické podklady pro tuto podkapitolu.

3.8.1 Schoenfeldova rezidua

Pro možnost využití Coxova semiparamatrického modelu nebo parametrického PH modelu je nutné, aby byla dodržena podmínka proporcionality hazardu. To znamená, že poměr rizik dvou subjektů je nezávislý na čase a jejich hazardní funkce jsou paralelní. Posouzení tohoto předpokladu je možné aplikací více přístupů, jednou z možností, pokud se jedná o Coxův model, je využití Schoenfeldových reziduí. Hodnota těchto reziduí je dána pro každý časový okamžik, ve kterém došlo k výskytu koncové události a je určena jako rozdíl mezi průměrnou hodnotou proměnné pro pozorování, které bylo ukončeno z důvodu výskytu sledované události a váženou průměrnou hodnotou této proměnné, tedy

$$\hat{r}_{ik} = c_i(x_{ik} - \hat{a}_{ik}), \tag{3.47}$$

kde i je jednotlivý subjekt, k značí dílčí vysvětlující proměnnou a

$$\hat{a}_{ik} = \frac{\sum_{j \in R(t_{(i)})} x_j \exp(x_j \beta)}{\sum_{j \in R(t_{(i)})} \exp(x_j \beta)}. \tag{3.48}$$

Hodnota těchto reziduí je pak porovnána s časem přežití (t), nebo některou z jeho modifikací, například $\ln(t)$ nebo Kaplan-Meierovým odhadem funkce přežití. První možností je výpočet koeficientu korelace mezi hodnotou reziduí a dobou přežití, druhou možností pak jsou grafické testy, kdy je zobrazena hodnota reziduí oproti době přežití. Vhodné je provést jak výpočet, tak grafické zobrazení, a to vůči co nejvíce modifikacím doby přežití. Přičemž aby byla splněna podmínka proporcionality hazardu, tak by neměla být zjištěna žádná závislost mezi pozorovanými veličinami a z grafického posouzení by neměl být patrný ani žádný výrazný trend.

3.8.2 Cox-Snellova rezidua

Hodnocení celkové adekvátnosti modelu lze provést pomocí Cox-Snellových reziduí. Tato rezidua jsou definována jako

$$r_{C_i} = \hat{H}_0(t_i)\exp(x_i\hat{\beta}), \tag{3.49}$$

kde odhady kumulativní hazardní funkce a parametru jsou získány z Coxova proporcionálního modelu. Odhad reziduí pro parametrický model s Weibullovým rozdělením pravděpodobnosti je upraven, neboť kumulativní hazardní funkce obsahuje i parametry měřítka a tvaru a hodnotu reziduí lze vypočítat pomocí vztahu

$$r_{C_i} = \hat{H}_i(t) = \exp(x_i\hat{\beta}) \cdot \hat{\lambda} t^{\hat{\lambda}}. \tag{3.50}$$

Odhadnutá Cox-Snellova rezidua jsou tedy vlastně odhadem kumulativní hazardní funkce příslušného modelu a jsou graficky porovnávána s kumulativní hazardní funkcí získanou pomocí neparametrických odhadů, tedy buď Kaplan-Meierovým odhadem (3.6), nebo Nelson-Aalenovým odhadem (3.8). Přičemž do odhadu funkce přežití pomocí těchto metod jsou dosazeny Cox-Snellova rezidua jako čas přežití. Pokud je model správný tak by měla Cox-Snellova rezidua odpovídat exponenciálnímu rozdělení pravděpodobnosti a kumulativní hazardní funkce by měla být přímka se sklonem 45°. Nelson-Aalenova kumulativní hazardní funkce s využitím reziduí by potom měla co nejvíce přiléhat této přímce, ovšem i v případě vhodného modelu je očekávána určitá variabilita kolem této linie, obzvláště na pravé straně grafu, a to z důvodu cenzorovaných dat a dřívějších výskytů koncových událostí, než by byla očekávána na základě sestaveného modelu. Vlastnosti reziduí je, že nejsou symetricky rozdělena kolem nuly a nemůžou být negativní.

3.8.3 Martingalova a deviantní rezidua

Martingalova a deviantní rezidua slouží k identifikování subjektů, u kterých se skutečná doba přežití výrazně liší oproti modelované době přežití a z tohoto důvodu mohou negativně zkreslovat odhadnuté parametry. Martingalova rezidua lze interpretovat jako rozdíl mezi skutečným počtem sledovaných událostí a počtem těchto událostí podle použitého modelu v daném časovém intervalu. Tato rezidua lze rovněž využít pro výběr správné formy proměnné, to znamená, že určitá proměnná by nemusela být do modelu zařazena ve své skutečné hodnotě, ale například v logaritmické nebo exponenciální úpravě.

Martingalova rezidua je možné využitím Cox-Snellových reziduí vyjádřit pro Coxův model i parametrický model s Weibullovým rozdělením vztahem

$$r_{M_i} = c_i - r_{C_i}. \tag{3.51}$$

Vzhledem k tomu, že c_i je časový cenzor, který nabývá hodnoty 0 pro cenzorovaná data a hodnoty 1 pro necenzorovaná data a Cox-Snellova rezidua nemohou být negativní, pak je zřejmé, že martingalova rezidua nabývají hodnot od mínus nekonečna do 1. Pokud reziduum pro daný subjekt je záporné a nabývá vyšších hodnot, pak u něj došlo k výskytu koncové události výrazně později, než předpovídá model. Oproti tomu pokud se hodnota rezidua blíží k jedné, pak k výskytu události došlo dříve, než by bylo určeno pomocí modelu. Dále je možné dokázat, že suma těchto reziduí je rovna 0 a v případě velkých souborů je i jejich očekávaná hodnota rovna 0.

Martingalova rezidua obdobně jako Cox-Snellova rezidua nejsou symetricky rozložena kolem nuly ani v případě, že model je adekvátní a tak z důvodu snadnější interpretace bývají více doporučována deviantní rezidua, které lze definovat jako

$$r_{D_i} = \operatorname{sgn}(r_{M_i})\left[-2(r_{M_i} + c_i \log(c_i - r_{M_i})\right]^{\frac{1}{2}}. \tag{3.52}$$

kde sgn je tzv. znaménková funkce, která nabývá hodnoty +1, když její argument je kladný a hodnoty -1, pokud je argument záporný. Tím je zajištěno, že deviantní rezidua mají stejné znaménko jako rezidua martingalova. Rozdíl mezi nimi pak je v oboru hodnot, který u deviantních reziduí je roven celé množině reálných čísel. Přičemž pokud je hodnota martingalových reziduí záporná a blíží se nekonečnu, pak se deviantní rezidua blíží nule. V případě, že se martingalova rezidua téměř rovnají svému maximum, tedy jedné, deviantní rezidua se přibližují nekonečnu.

4 Odhad modelů analýzy přežití

V praktické části této práce jsou aplikovány teoretické přístupy popsány v předešlých kapitolách na konkrétní data českých společností. V první části jsou analyzovány podniky napříč různými sektory pomocí neparametrických metod, což umožňuje vzájemné srovnání jednotlivých odvětví. V dalších částech je sestaven a verifikován semiparametrický Coxův model a parametrický model s proporcionálním hazardem a předpokladem, že doba přežití odpovídá Weibullovému rozdělení pravděpodobnosti. Tyto modely jsou sestaveny pouze pro sektor doprava. Výpočty a grafické zobrazení modelů byly získány využitím počítačového softwaru Stata 11 a doplňkové výpočty pomocí programu MS Excel.

4.1 Vstupní údaje

Pro potřeby analýzy přežití je zapotřebí počáteční a koncové datum a hodnoty nezávislých proměnných. Tyto údaje byly získány z informační databáze Magnusweb, případně dohledané ze serveru justice.cz. Do analýzy vstupují v této práci všechny české firmy působící ve vybraných odvětvích zapsané v obchodním rejstříku se vznikem od roku 1990 do roku 2005, za předpokladu, že byly v roce 2005 stále fungující a byly u nich zveřejněny výsledky hospodaření za tento rok. Za počáteční datum je v práci považováno datum vzniku společnosti podle obchodního rejstříku. Koncové datum představuje den bankrotu společnosti, v případě že k bankrotu došlo. Pokud k bankrotu nedošlo ve sledovaném období, tak je jako koncové datum zvoleno datum 31. 3. 2015, kdy došlo k ukončení pozorování. K tomuto datu ještě nejsou k dispozici výsledky hospodaření za rok 2014, hodnota vysvětlujících proměnných je proto převzata z finančních výkazů pro rok 2013. Pokud však u firmy nebyly známy finanční údaje za rok 2013, pak koncové datum představuje poslední den, za který byly finanční údaje známy.

Hodnoty vstupních proměnných byly vypočítány z finančních výkazů jednotlivých společností a celkem bylo spočítáno 22 poměrových ukazatelů z oblasti rentability, aktivity, likvidity, zadluženosti a stability. Výsledné hodnoty poměrových ukazatelů pro každou společnost byly získány aplikací vzorců (2.1) - (2.22). Sledována je rovněž velikost podniku, vyjádřena výší aktiv a přirozeným logaritmem celkových aktiv.

Analyzovány byly firmy, jejichž činnost působnosti je v jednom z devíti vybraných sektorů. Těmi jsou administrativa, doprava, informační technologie, kultura, stavebnictví, ubytování, zemědělství, zdravotní péče a zpracování vody a odpadu. Do těchto skupin byly zařazeny všechny české firmy, které vznikly od roku 1990 do roku 2005, za předpokladu, že byly v roce 2005 stále fungující a byly u nich zveřejněny výsledky hospodaření za tento rok. Celkem se jedná o 16 513 společností, z nichž u 1504 došlo ve sledovaném období k bankrotu. Přehled o počtu společností v daném odvětvi a počtu bankrotu je uveden v tabulce (Tabulka 4.1). Nejvíce zastoupeným sektorem je stavebnictví s 4548 firmami, kde také došlo k největšímu počtu bankrotujících podniků (669). Následuje odvětví IT a administrativa s více než dvěma tisíci podniky, počet bankrotů je však v těchto sektorech nižší než u ubytování a dopravy, kde byl tedy vyšší poměr mezi počtem bankrotů a počtem společností. Nejméně zastoupeny jsou sektory kultura, zdravotní péče a zpracování vody a odpadu, kde počet podniků je nižší než tisíc a počet bankrotů menší než padesát.

Tabulka 4.1 Přehled o vybraných odvětvích

Odvětví	Počet společností	Počet bankrotů
Administrativa	2052	127
Doprava	1229	232
Informační technologie	2629	96
Kultura	787	43
Stavebnictví	4548	669
Ubytování	1903	196
Zdravotní péče	646	12
Zemědělství	2051	96
Zpracování vody a odpadu	668	33
Celkem	**16513**	**1504**

Zdroj: magnusweb.bisnode.cz; vlastní zpracování

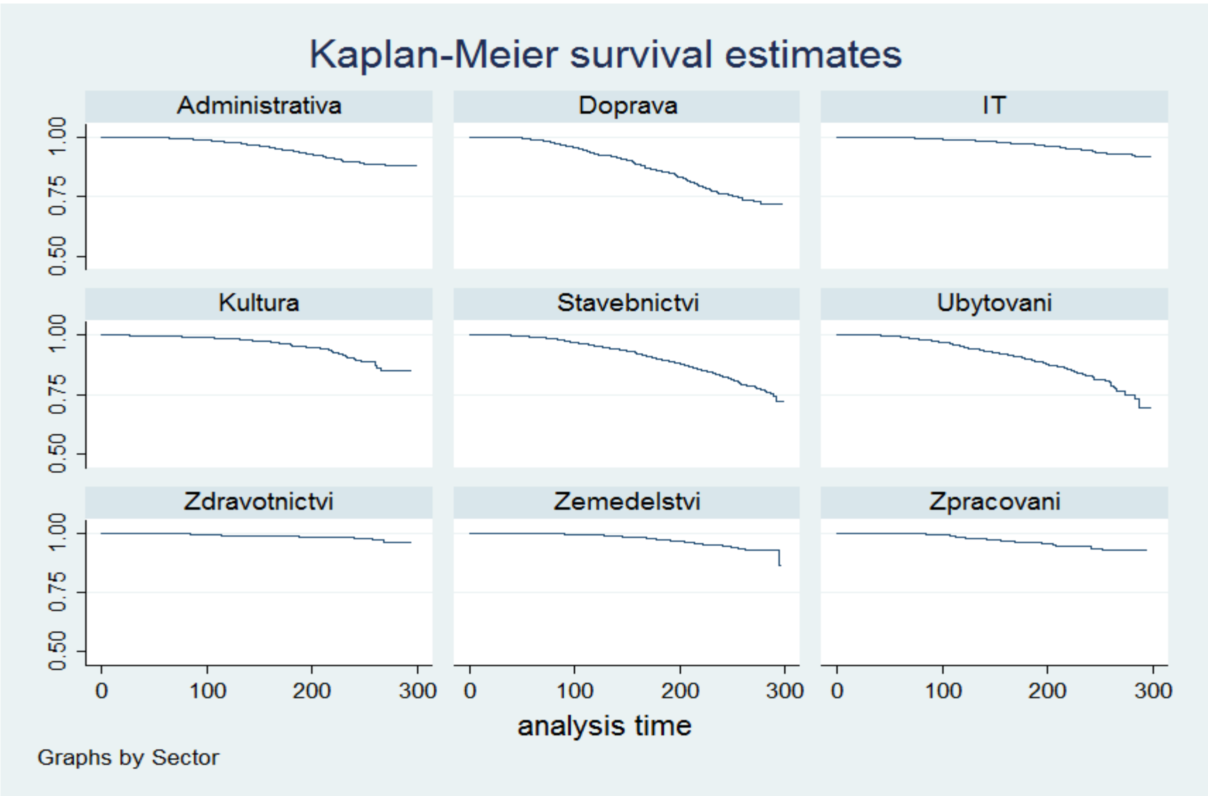
4.2 Neparametrické odhady

V této podkapitole jsou odhadnuty funkce přežití a hazardní funkce jednotlivých odvětví využitím neparametrických metod, tedy Kaplan-Meierova a Nelson-Aalenova odhadu, které jsou použity pro grafické znázornění funkce přežití a kumulativní hazardní funkce. Dále je proveden logrank test a zobecněný Wilcoxonův test pro statistické posouzení shodnosti funkce přežití každé dvojice sektorů.

Kaplan-Meierův odhad je vypočítán pomocí vzorce (3.6). Na obrázku (Obrázek 4.1) jsou znázorněny funkce přežití pro každý sektor jednotlivě. Na ose x je čas přežití v měsících, na ose y pravděpodobnost, že u společnosti nedojde k bankrotu do daného časového okamžiku. Bod na křivce je možné interpretovat jako míru pravděpodobnosti, s jakou u subjektu nedojde k bankrotu v daném čase od vzniku společnosti. To znamená, že například firma působící v dopravě má 80% pravděpodobnost, že nezbankrotuje po dvou stech měsících od svého vzniku. Funkce přežití u všech sektorů jsou klesající, což odpovídá předpokladům, neboť pravděpodobnost přežití by v čase měla klesat.

Na první pohled je zřejmá velká podobnost u sektoru doprava, stavebnictví a ubytování, kde pravděpodobnost dožití poměrně výrazně klesá. I u těchto křivek však je rozdíl, neboť zatímco u dopravy pokles s časem zpomaluje, u ubytování a stavebnictví naopak šance dožití klesá rychleji. U ubytování pak lze vidět v pozdějším období schodovitost funkce, což je typický znak funkce přežití, kdy už byl ve sledování menší počet subjektů a je tak patrný delší časový okamžik mezi jednotlivými bankroty. U firem z odvětví administrativy docházelo k největšímu počtu koncových událostí mezi stým a dvoustým měsícem, kdy křivka klesá významněji než na počátku nebo na konci období. Konstantní pokles je u sektoru informačních technologií (IT), kde pravděpodobnost přežití mají firmy vyšší než u předcházejících oborů. Zajímavý je pohled na funkci přežití pro odvětví kultura, jelikož zde mají podniky funkci přežití pomalu klesající až do přibližně dvou třetin grafu, pak ovšem klesá pravděpodobnost přežití rychleji. Funkce přežití zbývajících tří sektorů, tedy zdravotnictví, zemědělství a zpracování vody a odpadu, jsou si podobné a firmy z těchto odvětví mají největší pravděpodobnost, že u nich do 300 měsíců od založení nedojde k bankrotu. Tuto šanci mají největší podniky zabývající se zdravotní péčí. U zemědělských společností a podniků zabývajících se zpracováním vody a odpadu pravděpodobnost po určité době klesá více, ovšem i tak je pravděpodobnost bankrotu nízká.

Obrázek 4.1 Kaplan-Meierův odhad funkce přežití pro jednotlivá odvětví

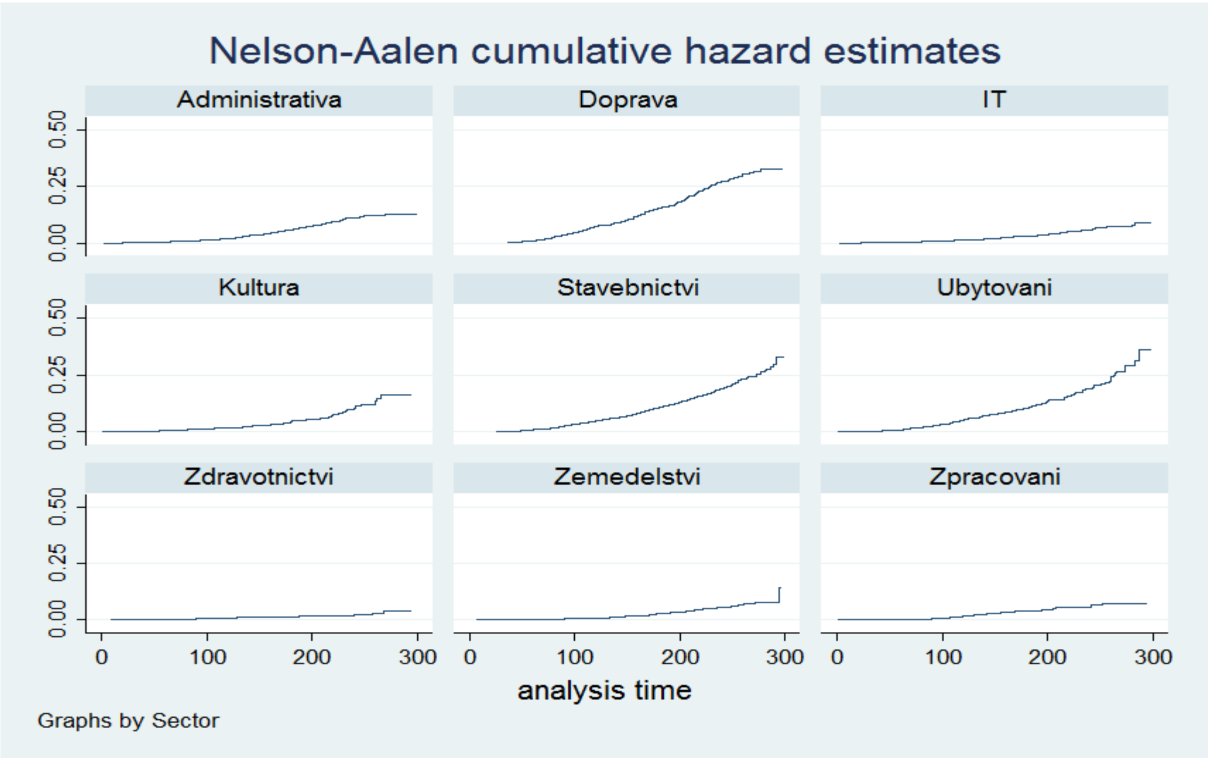


Zdroj: vlastní zpracování

Druhou základní funkcí v analýze přežití je hazardní funkce. Na obrázku (Obrázek 4.2) jsou znázorněny kumulativní hazardní funkce pro jednotlivá odvětví. Hodnoty použity pro zobrazení jsou vypočítány pomocí Nelson-Aalenova odhadu (3.7). Na ose x je opět doba v měsících, na ose y pravděpodobnost, že u subjektu dojde k bankrotu v daném časovém okamžiku. Tedy například podnik ze stavebního průmyslu má asi 15% šanci, že u něj dojde po 200 měsících k bankrotu, za předpokladu, že k bankrotu nedošlo do té doby. Kumulativní hazardní funkce je pro všechny odvětví rostoucí, protože riziko bankrotu v čase také stoupá.

Vzhledem k tomu, že vstupní data, tedy počet firem a počet bankrotů, jsou stejná pro výpočet hazardní funkce i pro výpočet funkce přežití, tak i srovnání kumulativních hazardních funkcí pro jednotlivá odvětví je podobné, jako porovnání funkcí přežití. Nejrychleji rostoucí je kumulativní hazardní funkce pro sektor doprava, ubytování a stavebnictví, přičemž u dopravy je tempo růstu stále pomalejší. Společnosti v těchto odvětvích mají tedy největší pravděpodobnost výskytu bankrotu. U administrativy je patrný největší růst rizika uprostřed sledované doby, pro firmy působící v informačních technologiích pak je zřejmý lineární nárůst pravděpodobnosti bankrotu. Společnosti podnikající v kulturní oblasti mají vyšší šanci bankrotu až v pozdější době. Nejmenší riziko výskytu koncové události je pro společnosti působící v odvětví zdravotnictví, nízké riziko je rovněž u zemědělských a zpracovatelských podniků.

Obrázek 4.2 Nelson-Aalenův odhad kumulativní hazardní funkce pro jednotlivá odvětví



Zdroj: vlastní zpracování

Po porovnání funkce přežití a kumulativní hazardní funkce pro dané sektory grafickými metodami je užitečné provést statistické testování shodnosti doby přežití napříč jednotlivými skupinami. Jak lze vidět v předcházejících grafech, u některých skupin je zřejmá vzájemná podobnost odhadnutých křivek. Pro posouzení statistické významnosti této podobnosti je použit logrank test a Wilcoxonův test. Testová statistika pro oba testy je vypočítána využitím rovnice (3.11), liší se pouze určení váhy. Pokud platí nulová hypotéza, pak jsou funkce přežití pro dané skupiny subjektů stejné. V tabulce (Tabulka 4.2) jsou všechny dvojice, u kterých p-hodnota pro alespoň jeden z testů byla vyšší než 0,05. To znamená, že všechny dvojice, které nejsou uvedeny v této tabulce, mají rozdílné funkce přežití na 95% hladině spolehlivosti. Naopak dvojice odvětví uvedené v tabulce nemají rozdílné funkce přežití na této hladině spolehlivosti, ale jejich funkce mají jistou podobnost. Detailní výsledky obou testů jsou uvedeny v přílohách č.1 a 2.

Na základě výsledků Wilcoxonova testu má osm dvojic podobnou funkci přežití, pomocí logrank testu se jedná o šest dvojic. Rozdíl způsobuje odvětví informačních technologií s kulturou a administrativa se zpracovatelským průmyslem. I u ostatních dvojic jsou zřetelné výrazné odlišnosti p-hodnoty mezi jednotlivými testy. Důvodem je větší váha Wilcoxonova testu na počátek sledovaného období, kdy je více subjektů podrobeno riziku, než na konec období. Ojedinělé výskyty koncových události na konci pozorování tak nemají takovou váhu jako u logrank testu, kdy váha je konstantní v průběhu celé doby. Nejvíce shodností s ostatními sektory má zpracovatelský průmysl, jehož funkce přežití je shodná na 5% hladině významnosti s funkcí přežití čtyř sektorů podle Wilcoxova testu, podle logrank testu se třemi. Shodnost funkcí přežití se třemi sektory má odvětví informačních technologií. Z osmi shodností tak v šesti případech je alespoň jedním ze sektorů zpracovatelský průmysl, nebo odvětví informačních technologií. Zbylé dvě dvojice tvoří sektor administrativy s odvětvím kultury a také odvětví stavebnictví s ubytováním. Podle obou testů nemá ani jednu podobnost dopravní sektor, jehož funkce přežití i kumulativní hazardní funkce se lišila už u grafického zobrazení. Liší se také zdravotnictví, kde je to však způsobeno především malým počtem bankrotů ve sledovaném období.

Tabulka 4.2 Výsledky logrank testu a Wilcoxonova testu

Odvětví		p-hodnota Wilcoxon	p-hodnota logrank
Administrativa	Kultura	**0,221**	**0,5343**
Administrativa	Zpracování	**0,0585**	**0,0153**
Kultura	IT	**0,0568**	**0,006**
Kultura	Zpracování	**0,526**	**0,1004**
IT	Zemědělství	**0,4527**	**0,6544**
IT	Zpracování	**0,2681**	**0,5569**
Zemědělství	Zpracování	**0,1471**	**0,3979**
Stavebnictví	Ubytování	**0,6673**	**0,623**

Zdroj: vlastní zpracování

4.3 Coxův proporciální hazardní model

Po neparametrických odhadech, kde jsou analyzovány funkce přežití a hazardní funkce pro různé sektory, je dále pracováno pouze se sektorem dopravy. Každý sektor má svá specifika a není vhodné sestavovat jeden model pro společnosti ze všech odvětví. Jednak je každý sektor zastoupen jiným počtem společností a také pro každý sektor může být významný jiný ukazatel a mohlo by dojít k nevhodnému odhadu výsledného modelu, jehož vypovídající hodnota by byla nízká. Bylo by možné sestavit model pro každé odvětví zvlášť, ale to by bylo časově i obsahově velmi náročné. Dále je tedy vybráno odvětví dopravy, které bylo využitím logrank testu a Wilcoxonova testu charakterizováno jako jedinečné, neboť funkce přežití není shodná s funkcí přežití žádného dalšího sektoru. Výhodou tohoto odvětví pro účely další analýzy je také velký počet firem (1229) a relativně vysoký počet bankrotů (232). V této časti je zkoumán vliv jednotlivých faktorů z finanční analýzy na pravděpodobnost přežití jednotlivých firem působících v dopravě využitím Coxova semiparametrického modelu.

4.3.1 Výběr proměnných do Coxova modelu

V práci je uvažováno s 24 finančními ukazateli, které mohou ovlivňovat pravděpodobnost bankrotu společnosti. Základní statistické údaje o těchto proměnných jsou uvedeny v příloze č.3. Užší výběr proměnných vstupujících do analýzy je proveden podle metodiky popsané v předchozí kapitole (3.7). Prvním krokem výběru proměnných je sestavení Coxova semiparametrického modelu s každým ukazatelem jednotlivě. Proměnné jsou statisticky testovány Waldovou statistikou (3.45). Pokud je daná proměnná statisticky významná alespoň na 20% hladině významnosti, pak je zařazena dále do modelu. V tabulce (Tabulka 4.3) jsou zobrazeny výsledky Waldova testu pro každou vysvětlující proměnnou, od nejvýznamnější po nejméně významnou. Kritická hodnota pro 20% hladinu významnosti je v absolutní hodnotě 1,28. V sloupci označeném z_{vyp} jsou vypočítané hodnoty Waldovy statistiky a v dalším sloupci příslušné p-hodnoty. Pokud je vypočítána hodnota v absolutní hodnotě vyšší než kritická, pak je daná veličina statisticky významná. Druhým vodítkem může být porovnání p-hodnoty s požadovanou hladinou významnosti, kdy pokud je p-hodnota menší, pak je příslušná proměnná významná. Na základě těchto kritérií je možné určit, že 10 ukazatelů je významných s 80% pravděpodobností a budou vstupovat do další fáze sestavování modelu. Jedná se o přirozený logaritmus celkových aktiv, podíl vlastního kapitálu na aktivech, celkovou zadluženost, stupeň krytí stálých aktiv, dobu obratu aktiv, pohledávek a závazků, dále podíl zásob na oběžných aktivech, rentabilitu nákladů a okamžitou likviditu.

Tabulka 4.3 Waldova statistika a příslušná p-hodnota pro jednorozměrné Coxovy modely

Proměnná	z_{vyp}	p-hodnota	Proměnná	z_{vyp}	p-hodnota
ln Celkových aktiv	**-8,45**	**0**	Čistý pracovní kapitál	0,55	0,58
Podíl VK na aktivech	**-3,77**	**0**	Úrokové krytí	-0,48	0,631
Celková zadluženost	**4,13**	**0**	Pohotová likvidita	-0,27	0,789
Stupeň krytí stálých aktiv	**-4,11**	**0**	Pohledávky na OA	0,26	0,793
DO aktiv	**6,18**	**0**	Běžná likvidita	-0,26	0,796
DO pohledávek	**5,61**	**0**	Finanční páka	-0,24	0,807
DO závazků	**6,14**	**0**	Celková aktiva	0	0,894
Zásoby na OA	**3,63**	**0**	ROCE	-0,12	0,905
ROC	**2,9**	**0,004**	ROE	-0,07	0,943
Okamžitá likvidita	**-1,79**	**0,074**	Obrátka aktiv	-0,07	0,947
Zadluženost vlastního kapitálu	-0,71	0,476	DO zásob	-0,06	0,952
ROA	-0,55	0,58	ROS	0,05	0,959

Zdroj: vlastní zpracování

Následující fází po určení významných parametrů v jednorozměrné analýze je sestavení modelu se všemi významnými proměnnými dohromady a následná eliminace těch proměnných, které se jeví jako nevýznamné na 10% hladině významnosti. Statistický software nenalezl řešení pro všech 10 proměnných, následně tedy byly z modelu odebrány ukazatele doby obratu aktiv a doby obratu pohledávek, jejichž koeficienty byly sice statisticky významné, ale jejich hodnota byla blízká nule. Tyto proměnné budou v pozdější fázi opět do modelu zařazeny a bude zjištěna jejich významnost.

Výsledné hodnoty prvotního výsledného modelu jsou v tabulce (Tabulka 4.4). Jsou zde uvedeny hodnoty odhadnutých koeficientů aplikací rovnice (3.20), jejich směrodatná chyba (3.27), intervaly spolehlivosti (3.28), vypočítaná Waldova statistika (3.45) a příslušná p-hodnota. Do výsledného modelu mohou být zařazeny pouze proměnné, které jsou statisticky významné na 10% hladině významnosti. Tuto podmínku podle p-hodnoty nesplňuje ukazatel okamžité likvidity a zásoby na oběžných aktivech. P-hodnota pro ukazatel okamžité likvidity je vyšší než p-hodnota pro ukazatel poměru zásob na oběžných aktivech a tak bude nejprve z modelu odebrán tento ukazatel. Pro potvrzení adekvátnosti vyřazení okamžité likvidity je využit i test poměru věrohodností funkce, a to modifikací vztahu (3.42). Porovnávána je hodnota log-parciální věrohodnostní funkce modelu, kde je obsažen ukazatel okamžité likvidity s hodnotou této funkce pro model, ve kterém tento ukazatel není. Testová statistika je rovna

$$G = 2 \cdot (-193,67 - (-194,0,22)) = 0,704, \tag{4.1}$$

a má chí kvadrát rozdělení s jedním stupněm volnosti. Přičemž kritická hodnota tohoto rozdělení pro 90% pravděpodobnost je 2,7. Vypočítaná statistika je tedy nižší než kritická hodnota a je možné přijmout nulovou hypotézu, že koeficient je nulový, takže ukazatel okamžité likvidity je z modelu odstraněn.

Tabulka 4.4 Model 1- odhadnuté hodnoty Coxova modelu

Proměnná	Koeficient	Směrodatná chyba	z_{vyp}	p-hodnota	95% interval spolehlivosti	
ln Celkových aktiv	-0,23093	0,09826	-2,35	0,019	-0,42352	-0,03834
Podíl VK na aktivech	0,10215	0,04987	2,05	0,041	0,00440	0,19989
Celková zadluženost	0,10954	0,05028	2,18	0,029	0,01099	0,20809
Stupeň krytí stálých aktiv	-0,01333	0,00347	-3,85	0,000	-0,02012	-0,00653
DO závazků	2,49E-06	4,05E-07	6,15	0,000	0,00000	0,00000
Zásoby na oběžných aktivech	8,78E-07	5,56E-0,7	1,58	0,114	0,00000	0,00000
ROC	0,00007	0,00003	2,65	0,008	0,00002	0,00012
Okamžitá likvidita	-0,00800	0,00973	-0,82	**0,411**	-0,02707	0,01107
Hodnota log-parciální věrohodnostní funkce		-193,67				

Zdroj: vlastní zpracování

Vyřazením ukazatele okamžité likvidity zůstalo 7 proměnných a byl odhadnut nový model. Odhadnuté hodnoty pro tento model jsou v tabulce (Tabulka 4.5). Opět je nutné porovnat p-hodnoty odhadnutých koeficientů se stanovenou hladinou významnosti ve výši 10 %. Jedinou proměnnou, jejíž p-hodnota je vyšší než kritických 0,1, je podíl zásob na oběžných aktivech. Tento ukazatel by tak měl být z modelu vyřazen, což je ještě vhodné potvrdit testem parciální věrohodnostní funkce. Testová statistika je nyní

$$G = 2 \cdot (-194,0,22 - (-201,24)) = 14,44. \tag{4.2}$$

V tomto případě je vypočítaná hodnota výrazně vyšší než kritická hodnota 2,7, nelze tak přijmout nulovou hypotézu, že ukazatel je statisticky nevýznamný. Došlo tak k rozporu mezi výsledky testování koeficientu pomocí Waldovy statistiky a testu věrohodnostní funkce a je otázkou, zda ukazatel vyloučit. Vzhledem k tomu, že hodnota ukazatele je téměř nulová, je tento ukazatel z modelu vyloučen, neboť by mohl ovlivňovat ostatní odhadnuté koeficienty a doba přežití subjektu není závislá na hodnotě této proměnné.

Tabulka 4.5 Model 2- odhadnuté hodnoty Coxova modelu

Proměnná	Koeficient	Směrodatná chyba	z_{vyp}	p-hodnota	95% interval spolehlivosti	
ln Celkových aktiv	-0,22714	0,09826	-2,31	0,021	-0,41971	-0,03456
Podíl VK na aktivech	0,10353	0,04985	2,08	0,038	0,00582	0,20124
Celková zadluženost	0,11104	0,05026	2,21	0,027	0,01252	0,20956
Stupeň krytí stálých aktiv	-0,01341	0,00346	-3,80	0,000	-0,02019	-0,00662
DO závazků	2,49E-06	4,04E-07	6,16	0,000	0,00000	0,00000
Zásoby na oběžných aktivech	8,92E-07	5,56E-07	1,60	**0,109**	0,00000	0,00000
ROC	0,00007	0,00002	2,90	0,004	0,00002	0,00012
Hodnota log-parciální věrohodnostní funkce		**-194,022**				

Zdroj: vlastní zpracování

Nyní jsou známy proměnné, které jednoznačně mají vliv na dobu přežití daného subjektu. Výsledky odhadnutého modelu, který zahrnuje šest vysvětlujících proměnných, jsou v tabulce (Tabulka 4.6).

Tabulka 4.6 Model 3- odhadnuté hodnoty Coxova modelu

Proměnná	Koeficient	Směrodatná chyba	z_{vyp}	p-hodnota	95% interval spolehlivosti	
ln Celkových aktiv	-0,28742	0,08362	-3,44	0,001	-0,45132	-0,12353
Podíl VK na aktivech	0,08663	0,04785	1,81	0,070	-0,00716	0,18042
Celková zadluženost	0,09296	0,04794	1,94	0,052	-0,00100	0,18692
Stupeň krytí stálých aktiv	-0,01312	0,00343	-3,83	0,000	-0,01984	-0,00641
DO závazků	2,52E-06	4,05E-07	6,22	0,000	1,72E-06	3,31E-06
ROC	0,00007	0,00002	2,80	0,005	0,00002	0,00012
Hodnota log-parciální věrohodnostní funkce		**-201,246**				

Zdroj: vlastní zpracování

Tento model nemůže být považován za konečný model, neboť znovu musí být prozkoumány ty proměnné, které v jednorozměrné analýze byly nevýznamné, protože nyní ve vícerozměrném modelu už by mohly být významné. Do stávajícího modelu tak jsou postupně přidávány jednotlivé proměnné, které byly určeny jako nevýznamné v jednorozměrné analýze. Tímto způsobem bylo odhadnuto 16 nových modelů, které obsahují ukazatele obsažené v modelu 3 plus jednu další proměnnou. Pomocí Waldovy statistiky je zjišťováno, jestli přidaná proměnná je nyní statisticky významná a měla by být do modelu zařazena. Výsledné hodnoty Waldovy statisiky a příslušné p-hodnoty jsou v tabulce (Tabulka 4.7). Při srovnání p-hodnot s kritickou hladinou 10 % je zřejmé, že statisticky významné jsou proměnné celková aktiva, rentabilita aktiv a čistý pracovní kapitál. Tyto ukazatele je potřeba podrobněji prozkoumat a určit, zda by měly být zařazeny do konečného modelu.

Tabulka 4.7 Waldova statistika a p-hodnota pro každou proměnnou, o jejímž přidání do modelu je uvažováno

Proměnná	z_{vyp}	p-hodnota	Proměnná	z_{vyp}	p-hodnota
Celková aktiva	1,74	**0,083**	Obrátka aktiv	-1,14	0,256
ROA	2,21	**0,027**	DO aktiv	1,18	0,236
ROE	-0,07	0,947	DO zásob	-0,17	0,867
ROS	0,11	0,908	DO pohledávek	-1,39	0,164
ROCE	-0,07	0,945	Běžná likvidita	-0,23	0,818
Finanční páka	-0,13	0,898	Pohotová likvidita	-0,26	0,791
Zadluženost VK	-0,55	0,585	Pohledávky na OA	0,14	0,886
Úrokové krytí	-0,38	0,702	Čistý pracovní kapitál	3,27	**0,001**

Zdroj: vlastní zpracování

První uvažovaným ukazatelem jsou celková aktiva. Na základě Waldovy statistiky by tento ukazatel měl být do modelu zařazen. Dalším kritériem je test poměru parciální log-věrohodnostní funkce. Testová statistika nabývá hodnoty

$$G = 2 \cdot (-200,6 - (-201,24)) = 1,28. \tag{4.3}$$

Referenční hodnotou je opět kritická hodnota chí kvadrát rozdělení pravděpodobnosti s jedním stupněm volnosti a pravděpodobnosti 90 %, což je 2,71. Z toho vyplývá, že není možné zamítnout nulovou hypotézu, že koeficient je nulový. Nebylo tak potvrzeno, že by ukazatel celkových aktiv měl být do modelu přidán. Dalším důvodem, který je proti zařazení, je skutečnostt, že v modelu již je obsažena tato proměnná, jen matematicky upravená, a to přirozený logaritmus celkových aktiv. Je tedy zbytečné, aby v modelu byly obě tyto proměnné.

Druhou proměnnou, která by na základě Waldovy statisitky měla být zařazena do modelu je rentabilita aktiv. Vypočítaná statistika testu parciální log-věrohodnostní funkce je nyní

$$G = 2 \cdot (-198,81 - (-201,24)) = 4,86. \tag{4.4}$$

Test potvrdil významnost ukazatele rentability aktiv a podle těchto kritérií by tato proměnná měla být do modelu zařazena. V dalším kroku byl proveden odhad nového modelu i s rentabilitou aktiv, výsledky jsou v tabulce (Tabulka 4.8). Po přidání této proměnné se významně zvýšila p-hodnota ukazatele podílu vlastního kapitálu na aktivech na 0,715. Tento nárůst je významný a takováto proměnná by v konečném modelu neměla být obsažena. V modelu tedy nemůžou být zároveň proměnné rentabilita aktiv a podíl vlastního kapitálu na aktivech. V prvotní analýze byl jako významný určen podíl vlastního kapitálu na aktivech, z tohoto důvodu bude tento ukazatel v modelu ponechán a nebude přidána proměnná rentabilita aktiv.

Tabulka 4.8 Odhadnuté hodnoty Coxova modelu po přidání ROA

Proměnná	Koeficient	Směrodatná chyba	z_{vyp}	p-hodnota	95% interval spolehlivosti	
ln Celkových aktiv	-0,25303	0,09028	-2,80	0,005	-0,42997	-0,07609
Podíl VK na aktivech	0,01913	0,05247	0,36	0,715	-0,08371	0,12197
Celková zadluženost	0,11854	0,05032	2,36	0,018	0,01991	0,21718
Stupeň krytí stálých aktiv	-0,01303	0,00346	-3,77	0,000	-0,01982	-0,00625
DO závazků	2,24E-06	4,77E-07	4,70	0,000	1,31E-06	3,17E-06
ROC	0,00006	0,00002	2,50	0,013	0,00001	0,00011
ROA	0,14762	0,06687	2,21	0,027	0,01656	0,27868
Hodnota log-parciální věrohodnostní funkce		**-201,246**				

Zdroj: vlastní zpracování

Poslední proměnnou, o jejímž zařazení do modelu na základě Waldovy statistiky je uvažováno, je čistý pracovní kapitál. Zjištěná statistika testu parciální log-věrohodnostní funkce je v tomto případě

$$G = 2 \cdot (-198,97 - (-201,24)) = 4,54. \qquad (4.5)$$

I v tomto případě je testová statistika vyšší než kritická hodnota a je zamítnuta hypotéza, že koeficient je nulový. Obdobně jako v předchozím případě u rentability aktiv i ukazatel čistého pracovního kapitálu zvýšil p-hodnotu podílu vlastního kapitálu na aktivech tak, že se tento ukazatel stal nevýznamný, jak lze vidět v tabulce (Tabulka 4.9). Navíc koeficient pro čistý pracovní kapitál je 2,13E-06, tedy téměř nulový a doba přežití pro jednotlivý subjekt tak není významně závislá na hodnotě tohoto ukazatele. Z tohoto důvodu ani čistý pracovní kapitál nebude přidán do výsledného modelu a za konečný model tak je označen model 3 (Tabulka 4.6).

Tabulka 4.9 Odhadnuté hodnoty Coxova modelu po přidání Čistého pracovního kapitálu

Proměnná	Koeficient	Směrodatná chyba	z_{vyp}	p-hodnota	95% interval spolehlivosti	
ln Celkových aktiv	-0,33059	0,08016	-4,12	0,000	-0,48771	-0,17348
Podíl VK na aktivech	0,07604	0,04736	1,61	0,108	-0,01679	0,16887
Celková zadluženost	0,08169	0,04736	1,72	0,085	-0,01113	0,17452
Stupeň krytí stálých aktiv	-0,01305	0,00342	-3,82	0,000	-0,01975	-0,00635
DO závazků	2,56E-06	4,06E-07	6,31	0,000	1,77E-06	3,36E-06
ROC	0,00007	0,00002	2,75	0,006	0,00002	0,00012
Čistý pracovní kapitál	2,13E-06	6,51E-07	3,27	0,001	8,54E-07	3,41E-06
Hodnota log-parciální věrohodnostní funkce		**-201,246**				

Zdroj: vlastní zpracování

Cílem této podkapitoly bylo určení statisticky významných proměnných, které ovlivňují dobu přežití jednotlivých společností. V prvním kroku byly identifkovány ukazatele, které jsou statisticky významné v jednorozměrné analýze. Další fázi bylo sestavení Coxova semiparametrického modelu s těmito proměnnými a eliminace těch proměnných, které byly nevýznamné. Po tomto procesu zůstalo v modelu 6 proměnných, jedná se o přirozený logaritmus celkových aktiv, stupeň krytí stálých aktiv, podíl vlastního kapitálu na aktivech, celkovou zadluženost, dobu obratu závazků a rentabilitu nákladů. Dále bylo uvažováno o přidání proměnných, které se v jednorozměrné analýze jevily jako nevýznamné, ale po bližším zkoumání nebyla žádná taková proměnná do modelu zahrnuta. Jako výsledný tak byl určen původní model s šesti výše uvedenými proměnnými, který je uveden v tabulce (Tabulka 4.6)

4.3.2 Interpretace výsledného Coxova modelu

Pro interpretaci výsledného modelu a jednotlivých ukazatelů je namísto odhadnutých koeficientů využíván hazardní poměr, který lze pro každou proměnnou získat aplikací vzorce (3.14), hodnoty koeficientů jsou vypočítány Newton-Raphsonovou iterační metodou pro rovnici (3.20), Hazardní poměr je získán exponenciální úpravou koeficientů, přičemž pokud je koeficient záporný, tak hazardní poměr je menší než 1 a pokud je kladný, tak je hazardní poměr vyšší než 1. Výsledné hodnoty odhadnutých koeficientů a hazardních poměrů jsou uvedeny v tabulce (Tabulka 4.10).

Tabulka 4.10 Odhadnuté koeficienty a hazardní poměry pro jednotlivé ukazatele

Proměnná	ln Celkových aktiv	Podíl VK na aktivech	Celková zadluženost	Stupeň krytí stálých aktiv	DO závazků	ROC
Koeficient	-0,28742	0,08663	0,09296	-0,01312	2,52E-06	0,00007
Hazardní poměr	0,75019	1,09050	1,09742	0,98696	1,000003	1,00007

Zdroj: vlastní zpracování

Pokud je hazardní poměr menší než 1, pak nárůst hodnoty proměnné vede ke snížení rizika bankrotu. Tento případ nastal u přirozeného logaritmu celkových aktiv a stupně krytí stálých aktiv. Hazardní poměr pro logaritmus aktiv je 0,75, což znamená, že při nárůstu přirozeného logaritmu aktiv o jednotku klesne pravděpodobnost bankrotu na 75 %. Dojde tedy ke snížení rizika o 25 %. To je z ekonomického hlediska pochopitelné, neboť čím větší firma, tím by měla být stabilnější a méně náchylná ke krachu. Stejně tak u ukazatele stupeň krytí stálých aktiv je žádoucí vyšší hodnota, protože hazardní poměr je ve výši 0,987 a tak nárůst ukazatele o jednotku vede ke snížení rizika bankrotu o 1,3 %. Tento ukazatel vyjadřuje podíl dlouhodobého kapitálu ke stálým aktivům, tudíž vyšší hodnota přispívá k vyšší stabilitě společnosti.

Hazardní poměr pro ostatní čtyři proměnné je vyšší než 1, nárůst hodnoty proměnné pak vede ke zvýšení pravděpodobnosti bankrotu. Ukazatel podílu vlastního kapitálu na aktivech nabývá hodnoty 1,0905, pokud tak dojde ke zvýšení tohoto ukazatele o jednotku, vzroste riziko bankrotu o 9,05 %. Tento výsledek je poměrně překvapující, vzhledem k tomu, že ukazatel vyjadřuje poměr vlastního kapitálu k celkovým aktivům a kvůli finanční stabilitě firmy bývá doporučována co nejvyšší hodnota tohoto ukazatele. Vysvětlení je možné nalézt v nákladech kapitálu, neboť náklady na vlastní kapitál jsou kvůli neomezené splatnosti a vyššímu riziku větší, než náklady na cizí kapitál. Pokud tedy podnik má vysokou hodnotu podílu vlastního kapitálu na aktivech, znamená to, že majetek firmy financují především vlastníci firmy a firma má vysoké náklady na kapitál, což může vést k finančním problémům ve společnosti. Podnik může přestat být schopen splácet své závazky a dostat se do platební neschopnosti. Malý podíl cizích zdrojů v podniku může být způsoben také neochotou věřitelů zapůjčovat peněžní prostředky danému podniku kvůli nedostatku důvěryhodnosti.

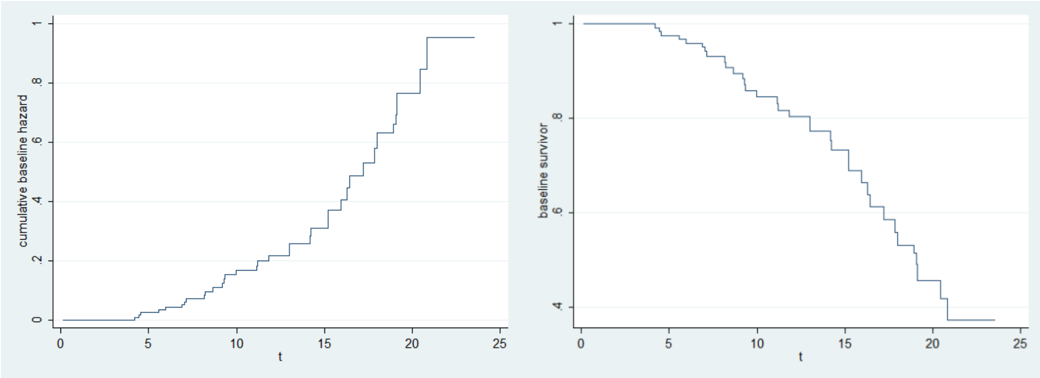
Odpovídající je hodnota hazardního poměru pro celkovou zadluženost, která je 1,0974, a zvýšení hodnoty o jednotku vede ke zvýšení rizika o 9,74 %. To odpovídá předpokladům, protože zvyšování zadluženosti firmy může vést k neschopnosti splácet své závazky a může vyústit v bankrot společnosti. Ukazatele doby obratu závazků a rentability nákladů jsou sice také vyšší než jedna, ovšem zvýšení těchto ukazatelů o jednotku má jen nepatrný vliv na zvýšení rizika selhání firmy, konkrétně o 0,0003 % a pro dobu obratu závazků a o 0,007 % pro rentabilitu nákladů.

Výsledný Coxův model je možné dosazením koeficientů do rovnice (3.13) zapsat ve tvaru

$$\begin{aligned} h(t,x,\beta) = h_0(t)\cdot \exp(&-0,2874\cdot \ln \text{ Celkových aktiv} + 0,0866\cdot \text{Podíl VK na aktivech} \\ &+0,09296\cdot \text{Celková zadluženost} - 0,01312\cdot \text{Stupeň krytí stálých aktiv} \\ &+0,000003\cdot \text{DO závazků} + 0,00007\cdot \text{ROC}). \end{aligned} \tag{4.6}$$

Coxův model neumožňuje určit přesně základní hazardní funkci, je však možné tuto funkci, stejně jako základní funkci přežití odhadnout. Odhadnuté funkce jsou zobrazeny na obrázku (Obrázek 4.3), v levé části je základní kumulativní hazardní funkce a v pravé části je základní funkce přežití. Tyto křivky jsou pro všechny subjekty shodné, neboť nezahrnují vliv jednotlivých proměnných, ale doba přežití je závislá pouze na čase. Vzhledem k tomu, že funkce přežití je odvozena od hazardní funkce přežití využitím rovnice (3.17), tak je zřejmá velká podobnost funkcí, když jsou navzájem inverzní. Pokud by pravděpodobnost přežití subjektu byla závislá pouze na době od vzniku dané společnosti, pak by bylo možné tvrdit, že pravděpodobnost bankrotu po přibližně 15 letech je asi 40 % a naopak, pravděpodobnost, že u podniku nedojde alespoň 15 let k výskytu bankrotu je 60%.

Obrázek 4.3 Odhad základní kumulativní hazardní funkce a základní funkce přežití



Zdroj: vlastní zpracování

4.4 Odhad parametrického modelu

Nevýhodou odhadnutého Coxova modelu je menší přesnost odhadnutých koeficientů, neboť není možné použit klasickou metodu maximální věrohodnosti. Aby bylo možné tuto metodu použít, tak musí být známo rozdělení pravděpodobnosti, kterým se řídí náhodná veličina. Dalším krokem tak bude sestavení parametrického modelu s proporcilnálním hazardem (PH model). To znamená, že odhadnuté hazardní poměry pro dané ukazatele jsou v čase konstantní a nejsou tedy na čase závislé. Stejná podmínka platí také pro Coxův semiparametrický model, a tyto modely tedy mohou být na závěr porovnány. Nutnou podmínkou pro sestavení parametrického modelu je určení, kterému rozdělení pravděpodobnosti odpovídá náhodná složka, tedy doba přežití. Běžně používaným rozdělením pravděpodobnosti je Weibullovo rozdělení, se kterým je pracováno i v této práci.

4.4.1 Výběr proměnných do parametrického modelu.

Výběr významných proměnných, které ovlivňují dobu přežití, je stejný jako u Coxova modelu. Rozdíl je pouze v testování pomocí log-věrohodnostní funkce, protože u Coxova modelu se jednalo o parciální věrohodnostní funkci, zatímco u parametrického modelu se jedná již o plnohodnotnou věrohodnostní funkci. Nejprve je nutné opět sestavit jednorozměrný model pro každou veličinu zvlášť. V tabulce (Tabulka 4.11) jsou uvedeny jednotlivé hodnoty Waldovy statistiky (3.45) a příslušné p-hodnoty. Kritická hladina významnosti je 20 % a tak je možné určit, že významných proměnných je celkem 10, přičemž se jedná o stejné ukazatele jako v semiparametrickém modelu.

Tabulka 4.11 Waldova statistika a příslušná p-hodnota pro jednorozměrné parametrické modely

Proměnná	z_{vyp}	p-hodnota	Proměnná	z_{vyp}	p-hodnota
ln Celkových aktiv	**-8,99**	**0**	ROA	-0,56	0,578
Podíl VK na aktivech	**-3,96**	**0**	Úrokové krytí	-0,48	0,634
Celková zadluženost	**4,35**	**0**	Pohotová likvidita	-0,26	0,796
Stupeň krytí stálých aktiv	**-4,06**	**0**	Běžná likvidita	-0,24	0,808
DO aktiv	**6,47**	**0**	Pohledávky na OA	0,23	0,819
DO pohledávek	**5,85**	**0**	Finanční páka	-0,2	0,839
DO závazků	**6,42**	**0**	ROCE	-0,17	0,868
Zásoby na OA	**4,86**	**0**	Celková aktiva	0,09	0,928
ROC	**2,87**	**0,004**	ROE	-0,06	0,951
Okamžitá likvidita	**-1,67**	**0,095**	Obrátka aktiv	-0,06	0,951
Zadluženost vlastního kapitálu	-0,65	0,517	DO zásob	-0,06	0,954
Čistý pracovní kapitál	0,56	0,572	ROS	0,03	0,975

Zdroj: vlastní zpracování

V další fázi jsou odstraňovány ty proměnné, které se po sestavení celkového modelu jeví být jako statisticky nevýznamné a to na základě Waldova testu a testu log-věrohodnostní funkce. Tento proces byl detailněji popsán u Coxova modelu, a tak jsou nyní v tabulce (Tabulka 4.12) zobrazeny výsledné hodnoty pro předběžný model, kde už zůstaly pouze významné veličiny. Statisticky významné se ukázaly stejné proměnné jako u Coxova modelu, tedy přirozený logratimus celkových aktiv, podíl vlastního kapitálu na aktivech, celková zadluženost, stupeň krytí stálých aktiv, doba obratu závazků a rentabilita nákladů. Mimo tyto ukazatele jsou v tabulce rovněž údaje o odhadnuté konstantě, ze které je možné vypočítat parametr lambda, a parametru *p*, které jsou součásti parametrického modelu s Weibullovým rozdělením pravděpodobnosti. Tyto hodnoty jsou popsány v interpretaci výsledného modelu. Hodnoty koeficientů jsou získány Newton-Raphsonovou iterační metodou a využitím rovnice (3.37), směrodatná chyba (3.27) a intervaly spolehlivosti pomocí vztahu (3.38) až (3.40).

Tabulka 4.12 Odhadnuté hodnoty pro parametrický model

Proměnná	Koeficient	Směrodatná chyba	z_{vyp}	p-hodnota	95% interval spolehlivosti	
ln Celkových aktiv	-0,30663	0,08705	-3,52	0,000	-0,47724	-0,13601
Podíl VK na aktivech	0,10602	0,03767	2,81	0,005	0,03219	0,17986
Celková zadluženost	0,11237	0,03764	2,99	0,003	0,03860	0,18613
Stupeň krytí stálých aktiv	-0,01148	0,00325	-3,53	0,000	-0,01786	-0,00511
DO závazků	2,60E-06	3,94E-07	6,60	0,000	1,83E-06	3,37E-06
ROC	0,00007	0,00002	2,71	0,007	0,00002	0,00011
Konstanta	-23,17407	3,28153	-7,06	0,000	-29,60575	-16,74239
Parametr *p*	2,593068	0,368967			1,961989	3,427135
Hodnota log-věrohodnostní funkce		-129,42				

Zdroj: vlastní zpracování

Do tohoto modelu byly postupně přidávány jednotlivé proměnné, které byly v prvním kroku identifikovány jako statisticky nevýznamné. Po přidání do předběžného modelu se jako významné ukázaly tři proměnné, a to celková aktiva, rentabilita aktiv a ukazatel čistého pracovního kapitálu, jak je vidět z tabulky (Tabulka 4.13). U těchto proměnných je potřeba zvážit jejich přidání do výsledného modelu.

Tabulka 4.13 Waldova statistika a p-hodnota pro proměnné přidané do parametrického modelu

Proměnná	z_{vyp}	p-hodnota	Proměnná	z_{vyp}	p-hodnota
Celková aktiva	1,77	**0,076**	Obrátka aktiv	-1,11	0,266
ROA	1,99	**0,047**	DO aktiv	1,09	0,276
ROE	-0,06	0,949	DO zásob	-0,14	0,89
ROS	0,12	0,904	DO pohledávek	-1,58	0,115
ROCE	-0,07	0,946	Běžná likvidita	-0,23	0,817
Finanční páka	-0,13	0,65	Pohotová likvidita	-0,26	0,792
Zadluženost VK	-0,45	0,893	Pohledávky na OA	0,19	0,851
Úrokové krytí	-0,41	0,68	Čistý pracovní kapitál	3,39	**0,001**

Zdroj: vlastní zpracování

První uvažovanou proměnnou, která na základě Waldovy statistiky je významná po přidání do modelu jsou celková aktiva. Odhadnutý model po přidání celkových aktiv je v tabulce (Tabulka 4.14). Pro potvrzení významnosti je proveden test log-věrohodnostní funkce. Vypočítána G statistika je v tomto případě

$$G = 2 \cdot (-128,75 - (-129,24)) = 0,98. \qquad (4.7)$$

Do modelu jsou přidávány v této fázi proměnné, které jsou významné na 10% hladině významnosti, a tak je vypočítaná hodnota porovnávána s kritickou hodnotou chí kvadrát rozdělení s jedním stupněm významnosti a pravděpodobností 90 %, což je 2,71. Testová statistika je nižší než kritická hodnota, lze tedy na dané hladině významnosti přijmout nulovou hypotézu, že koeficient je nulový. Navíc už v původním modelu je ukazatel přirozeného logaritmu aktiv a celková aktiva tedy nejsou zařazena do výsledného modelu.

Tabulka 4.14 Hodnoty parametrického modelu po přidání celkových aktiv

Proměnná	Koeficient	Směrodatná chyba	z_{vyp}	p-hodnota	95% interval spolehlivosti	
ln Celkových aktiv	-0,32415	0,085415	-3,8	0,000	-0,49156	-0,15674
Podíl VK na aktivech	0,101856	0,037396	2,72	0,006	0,02856	0,175151
Celková zadluženost	0,107871	0,0373	2,89	0,004	0,034763	0,180978
Stupeň krytí stálých aktiv	-0,01139	0,003249	-3,51	0,000	-0,01776	-0,00503
DO závazků	2,61E-06	3,94E-07	6,62	0,000	1,84E-06	3,38E-06
ROC	6,59E-05	2,46E-05	2,68	0,007	1,77E-05	0,000114
Celková aktiva	4,90E-08	2,77E-08	1,77	0,076	-5,17E-09	1,03E-07
Konstanta	-22,9815	3,278962	-7,01	0,000	-29,4082	-16,5549
Parametr *p*	2,588617	0,368294			1,958681	3,421149
Hodnota log-věrohodnostní funkce		-128,75				

Zdroj: vlastní zpracování

Rentabilita aktiv je další ukazatel, který je významný na základě Waldovy statistiky. Sestavený parametrický model s tímto ukazatelem je shrnut v tabulce (Tabulka 4.15). Vypočítaná statistika pro test log-věrohodnostní funkce je nyní

$$G = 2 \cdot (-127,27 - (-129,24)) = 3,94. \tag{4.8}$$

Tato hodnota je vyšší než kritická hodnota 2,71 a je nutné zamítnout nulovou hypotézu. Z toho vyplývá, že na základě obou testů by proměnná měla být do modelu přidána. Po přidání této proměnné však výrazně vzroste p-hodnota pro ukazatel podílu vlastního kapitálu na aktivech a to až na 0,329. Tento ukazatel však byl významný v dřívějších krocích a bude tak v modelu ponechán a naopak nebude přidán ukazatel rentability aktiv, neboť by negativně ovlivňoval zbylé proměnné.

Tabulka 4.15 Hodnoty parametrického modelu po přidání rentability aktiv

Proměnná	Koeficient	Směrodatná chyba	z_{vyp}	p-hodnota	95% interval spolehlivosti	
ln Celkových aktiv	-0,26713	0,095536	-2,8	0,005	-0,45438	-0,07988
Podíl VK na aktivech	0,042307	0,043342	0,98	0,329	-0,04264	0,127256
Celková zadluženost	0,137559	0,041056	3,35	0,001	0,057091	0,218027
Stupeň krytí stálých aktiv	-0,01073	0,003291	-3,26	0,001	-0,01718	-0,00428
DO závazků	2,36E-06	4,56E-07	5,17	0,000	1,46E-06	3,25E-06
ROC	5,94E-05	2,46E-05	2,41	0,016	1,11E-05	0,000108
ROA	0,142047	0,071392	1,99	0,047	0,00212	0,281973
Konstanta	-23,9962	3,365975	-7,13	0,000	-30,5934	-17,399
Parametr *p*	2,641478	0,378806			1,994246	3,498769
Hodnota log-věrohodnostní funkce	-127,27					

Zdroj: vlastní zpracování

Poslední proměnnou, která by mohla být přidána do modelu na základě Waldovy statistiky, je čistý pracovní kapitál. Výsledné hodnoty pro model obsahující tuto proměnnou jsou v tabulce (Tabulka 4.16). Zjištěná testová statistika pro test log-věrohodnostní funkce je rovna

$$G = 2 \cdot (-127,04 - (-129,24)) = 4,4. \tag{4.9}$$

Po srovnání této hodnoty s kritickou hodnotou chí kvadrát rozdělení s jedním stupněm volnosti a hladinou významnosti 10 % není možné přijmout nulovou hypotézu. Proměnná tak je statisticky významná na základě obou testů. Problémem však je, že i když statisticky je koeficient pro čistý pracovní kapitál nenulový, tak při pohledu na koeficient, který je ve výši 2,19E-06 je zřejmé, že koeficient je téměř nulový a doba přežití pro subjekt není výrazně závislá na tomto ukazateli. Čistý pracovní kapitál tedy také není přidán do předběžného modelu, neboť jeho vliv na nezávislou veličinu je nulový a mohl by akorát negativně zkreslovat hodnoty jiných proměnných.

Tabulka 4.16 Hodnoty parametrického modelu po přidání čistého pracovního kapitálu

Proměnná	Koeficient	Směrodatná chyba	z_{vyp}	p-hodnota	95% interval spolehlivosti	
ln Celkových aktiv	-0,35284	0,08255	-4,27	0	-0,51463	-0,19104
Podíl VK na aktivech	0,095344	0,036968	2,58	0,01	0,022889	0,167799
Celková zadluženost	0,100949	0,036766	2,75	0,006	0,028889	0,173008
Stupeň krytí stálých aktiv	-0,01134	0,003237	-3,5	0	-0,01768	-0,00499
DO závazků	2,65E-06	3,95E-07	6,7	0	1,87E-06	3,42E-06
ROC	6,51E-05	2,45E-05	2,65	0,008	0,000017	0,000113
Čistý pracovní kapitál	2,19E-06	6,48E-07	3,39	0,001	9,25E-07	3,46E-06
Konstanta	-22,7184	3,28135	-6,92	0	-29,1497	-16,2871
Parametr *p*	2,58554	0,368163			1,955898	3,417877
Hodnota log-věrohodnostní funkce		-127,04				

Zdroj: vlastní zpracování

V této části byly určené statisticky významné proměnné, které ovlivňují dobu přežití jednotlivých podniků. V prvním fázi byly identifkovány ukazatele, které jsou statisticky významné v jednorozměrné analýze. Dalším krokem bylo sestavení parametrického PH modelu za předpokladu, že doba přežití odpovídá Weibulovu rozdělení pravděpodobnosti, s těmito proměnnými a odebrání těch proměnných, které byly nevýznamné. Po tomto procesu zůstalo v modelu 6 proměnných, jedná se o přirozený logaritmus celkových aktiv, stupeň krytí stálých aktiv, podíl vlastního kapitálu na aktivech, celkovou zadluženost, dobu obratu závazků a rentabilitu nákladů. Dále bylo uvažováno o přidání proměnných, které se v jednorozměrné analýze jevily jako nevýznamné, ale po bližším zkoumání nebyla žádná taková proměnná do modelu zahrnuta a jako výsledný model byl určen ten, jehož hodnoty jsou v tabulce (Tabulka 4.12). Součástí tohoto modelu je také odhad parametru *p* a konstanty, která v další části slouží k odhadnutí parametru lambda.

4.4.2 Interpretace výsledného parametrického modelu

Interpretace zjištěných koeficientů pro jednotlivé proměnné je podobná jako u Coxova modelu, opět je využíván hazardní poměr (3.14). Hodnoty odhadnutých koeficientů jsou převzaty z tabulky (Tabulka 4.12), kde jsou odhadnuté koeficienty pro výsledný model. V tabulce (Tabulka 4.17) jsou zobrazeny hodnoty odhadnutých koeficientů a hazardní poměry pro jednotlivé ukazatele. Hazardní poměr pro přirozený logaritmus celkových aktiv a stupeň krytí stálých aktiv nabývá hodnoty menší než 1 a tak zvýšení těchto ukazatelů vede ke snížení pravděpodobnosti bankrotu. V případě logaritmu celkových aktiv způsobí zvýšení ukazatele o jednotku pokles rizika bankrotu na 73,59 %, tedy o 26,41 %. Nárůst proměnné stupeň krytí stálých aktiv o jednotku by vedlo k snížení rizika na 98,86 %, to znamená, že riziko by bylo nižší o 1,14 %.

Hazardní poměr pro další čtyři ukazatele je vyšší než jedna, tím pádem nárůst ukazatele má za následek zvýšení pravděpodobnosti bankrotu. Nejvyšší hodnota hazardní poměru je u celkové zadluženosti, kdy jednotkové zvýšení této veličiny by vedlo ke zvýšení rizika bankrotu o 11,89 %. Podobně i zvýšení ukazatele podílu vlastního kapitálu na aktivech o jednotku způsobí nárůst pravděpodobnosti výskytu úpadku o 11,19 %. Hodnoty hazardního poměru pro dobu obratu závazků a rentability nákladů jsou blízké jedné a mají jen nepatrný vliv na dobu přežití subjektů. Jednotkový nárůst jejich hodnot způsobí zvýšení pravděpodobnosti bankrotu o 0,0003 % u doby obraty závazků a o 0,007 % u rentability nákladů.

Tabulka 4.17 Odhadnuté koeficienty a hazardní poměry pro jednotlivé ukazatele v parametrickém modelu

Proměnná	ln Celkových aktiv	Podíl VK na aktivech	Celková zadluženost	Stupeň krytí stálých aktiv	DO závazků	ROC
Koeficient	-0,30663	0,10602	0,11237	-0,01148	2,60E-06	0,00007
Hazardní poměr	0,73592	1,11185	1,11892	0,98858	1,000003	1,00007

Zdroj: vlastní zpracování

Na rozdíl od Coxova modelu je v parametrickém modelu předpokládáno, že se doba přežití řídí daným rozdělením pravděpodobnosti, v tomto případě Weibullovým, a tak jsou součástí výsledného modelu i odhadnuté hodnoty pro konstantu, parametr měřítka lambda (3.30) a parametr p. Výsledné hodnoty pro tyto veličiny jsou v tabulce (Tabulka 4.18). Parametr p je pro všechny subjekty stejný a určuje tvar hazardní funkce. V tomto modelu je jeho hodnota 2,593, což je vyšší než jedna a tak je možné potvrdit, že hazardní funkce je rostoucí, a to je v souladu s očekáváním, neboť pravděpodobnost bankrotu subjektu by měla v čase růst.

Tabulka 4.18 Hodnoty pro jednotlivé parametry parametrického modelu

Parametr	Konstanta	lambda	*p*
Hodnota	-23,174	8,62E-11	2,593

Zdroj: vlastní zpracování

Z odhadnutých hodnot parametrů p a parametru lambda lze dosazením do rovnice (3.29) získat základní hazardní funkci, která je pro všechny subjekty stejná a je ve tvaru

$$h_0(t) = \lambda \cdot p \cdot t^{p-1} = 8,62E - 11 \cdot 2,593 \cdot t^{1,593} = -1,998E - 10 \cdot t^{1,593}. \qquad (4.10)$$

Tato funkce je tedy pro všechny subjekty totožná a dosazením časového údaje ve dnech namísto t lze zjistit pravděpodobnost úpadku k danému okamžiku, a to bez vlivu vysvětlujících proměnných. Výsledný model se základní hazardní funkcí a vysvětlujícími veličinami má dosazením do rovnice (3.31) následující tvar,

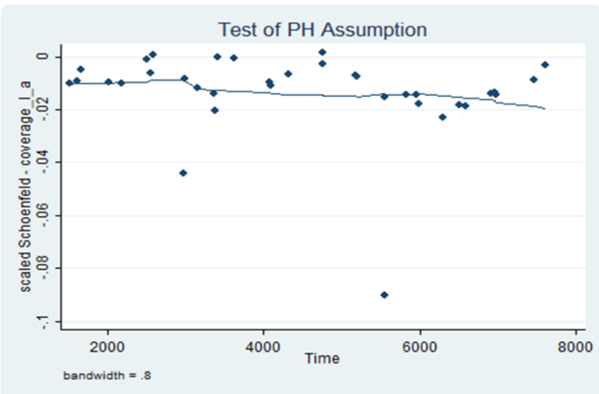
$$
\begin{aligned}
h(t,x,\beta,\lambda,p) = -1{,}998E-10\cdot t^{1{,}593}\cdot \exp(&-0{,}3066\cdot \ln\ \text{Celkových aktiv}\\
&+0{,}106\cdot \text{Podíl VK na aktivech} + 0{,}1124\cdot \text{Celková zadluženost}\\
&-0{,}0115\cdot \text{Stupeň krytí stálých aktiv}\\
&+0{,}000003\cdot \text{DO závazků} + 0{,}00007\cdot \text{ROC}).
\end{aligned} \qquad (4.11)
$$

4.5 Verifikace odhadnutých modelů

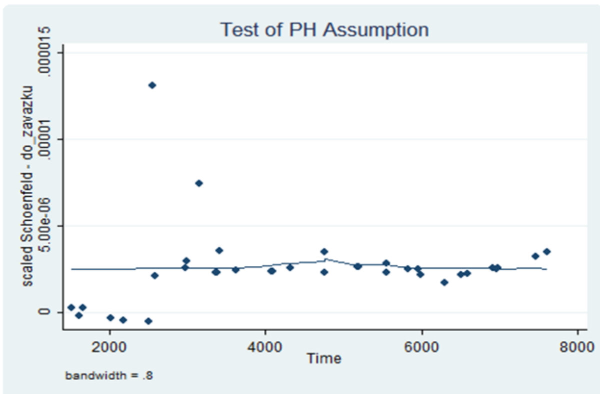
Nezbytným krokem po sestavení modelu je jeho verifikace. Nejdůležitějším předpokladem pro aplikaci Coxova modelu a vybraného parametrického modelu je dodržení proporcionality hazardu, tedy poměr rizik dvou subjektů je nezávislý na čase. Testování tohoto předpokladu je provedeno pro každou proměnnou v modelu zvlášť a je znázorněno na obrázku (Obrázek 4.4). Zde je zobrazeno 6 grafů, každý pro jednu proměnnou, na ose x je čas ve dnech, na ose y jsou zachycena Schoenfeldova rezidua vypočítána pomocí rovnice (3.47). Aby byla splněna podmínka proporcionality hazardu, tak by neměl být patrný žádný trend. Z výsledných grafů je možné potvrdit, že ani u jedné proměnné nelze určit trendovou linii, tudíž je splněna podmínka proporcionality hazardu pro každou proměnnou a není nutnost dalších úprav.

Obrázek 4.4 Test proporcionality hazardu využitím Schoenfeldových reziduí pro jednotlivé proměnné

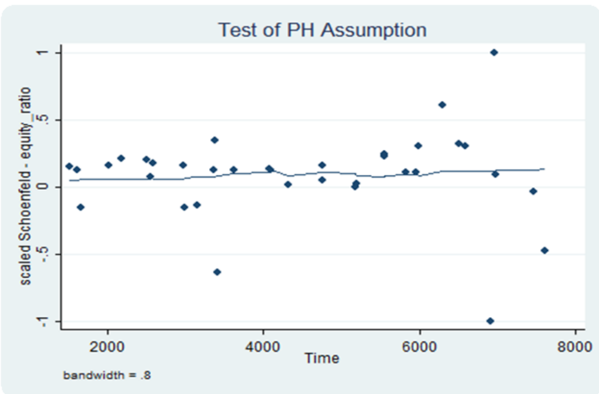
a, Stupeň krytí stálých aktiv



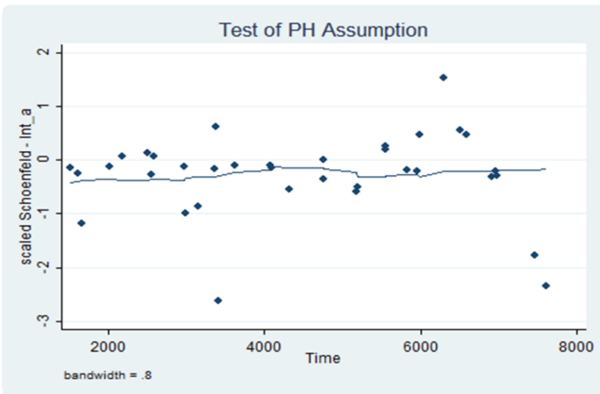
b, Doba obratu závazků



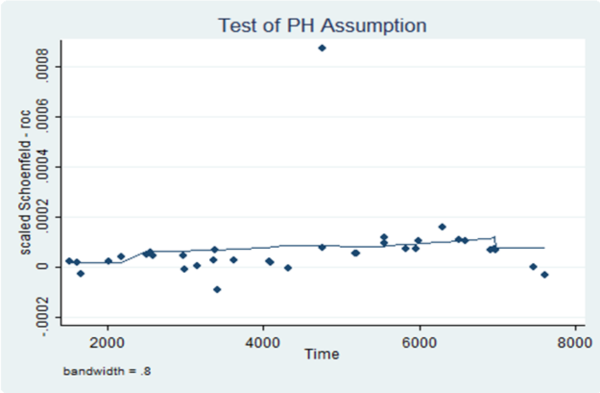
c, Podíl vlastního kapitálu na aktivech



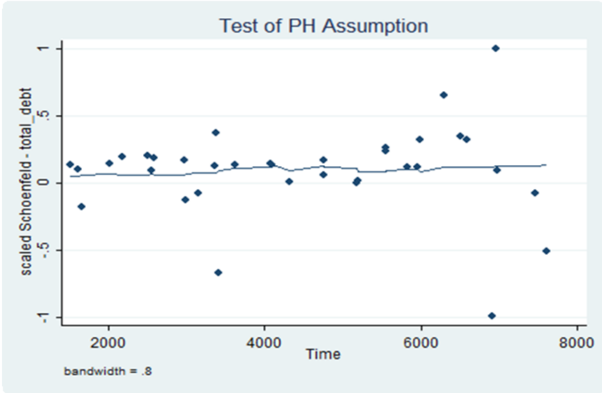
d, ln Celkových aktiv



e, Rentabilita aktiv



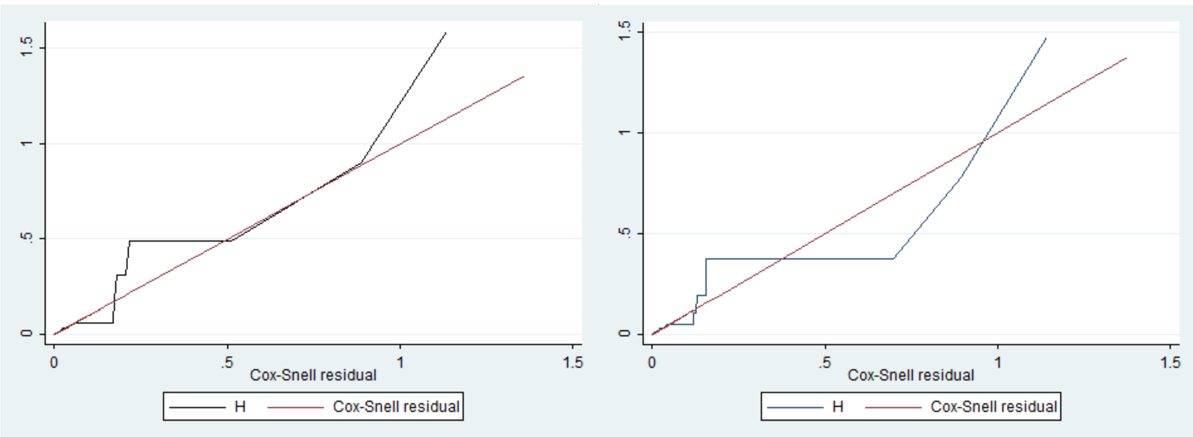
f, Celková zadluženost



Zdroj: vlastní zpracování

Pro ověření celkové adekvátnosti mohou sloužit Cox-Snellova rezidua. Na obrázku (Obrázek 4.5) je zobrazena referenční linie a kumulativní hazardní funkce pro tato rezidua, vypočítána využitím rovnice (3.49) pro Coxův model, a rovnice (3.50) pro parametrický model. V levém grafu jsou znázorněny tyto křivky pro Coxův model, v pravém pak pro parametrický model. V ideálním případě by kumulativní hazardní funkce a referenční linie měly být totožné. Velkou podobnost však lze vidět pouze v prostřední části grafu pro Coxův model, v levé i v pravé části grafu je zřejmá odchylka od linie 45 º. U parametrického modelu je patrné značné odchýlení v celé šíři grafu. Na druhou stranu na začátku a na konci grafu jsou odchylky u parametrického modelu menší než u Coxova modelu. Je tak obtížné určit, který z modelů vystihuje skutečnost lépe, jelikož uprostřed je to Coxův model, ale ve zbývajících částech to je model parametrický.

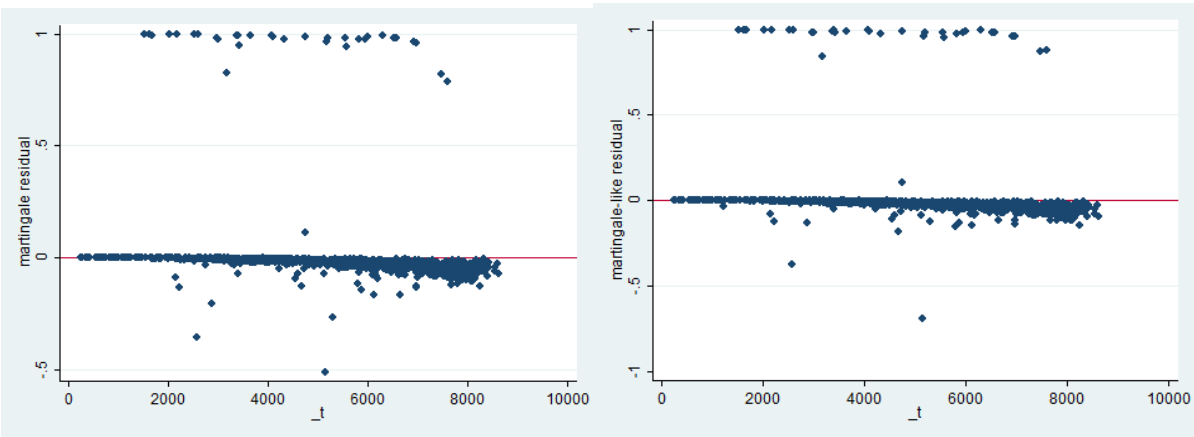
Obrázek 4.5 Cox-Snellova rezidua pro Coxův model a parametrický model



Zdroj: vlastní zpracování

Pro identifikaci subjektů, jejichž predikovaná doba přežití se výrazně liší od skutečné doby přežití, jsou využívána martingalova a deviantní rezidua. Martingalova rezidua zobrazena v grafu jsou vypočítána využitím vzorce (3.51) a nabývají hodnot $(-\infty; +1>$. Přičemž vyšší záporná hodnota značí, že u subjektu došlo k výskytu koncové události později, než by bylo určeno pomocí výsledného modelu. To se v tomto případě týká pouze pár subjektů. Ve srovnání se zápornými hodnotami je pro víc společností hodnota rezidua blízká 1, a to značí, že k bankrotu došlo výrazně dříve, než je modelováno. Hodnota rezidua pro většinu subjektů se však pohybuje kolem nuly, což značí adekvátnost modelu pro tyto subjekty. Z tohoto pohledu jsou si oba modely podobné, i když nejvyšší záporná hodnota rezidua pro subjekt je u parametrického modelu, kdy u jedné společnosti došlo k bankrotu výrazně později, než je předpokládáno.

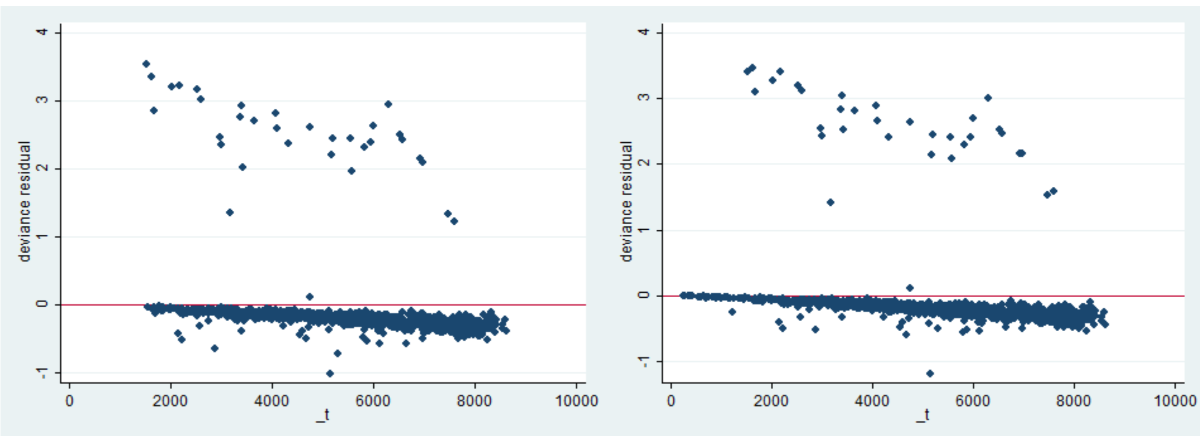
Graf 4.1 Martingalova rezidua pro Coxův model a parametrický model



Zdroj: vlastní zpracování

Obdobné zhodnocení jako martingalova rezidua lze provést pro deviantní rezidua zobrazena v grafu (Graf 4.2), vypočítána využitím rovnice (3.52). Subjekty, pro něž byla hodnota martingalových reziduí blízká 1, nyní mají hodnotu deviantních reziduí kladnou a v intervalu od 1 do 4. To jsou tedy podniky, u nichž nastal bankrot dříve než by bylo určeno pomocí modelu. Opět je viditelná velká podobnost a nelze jednoznačně určit, který model je vhodnější. Největší část reziduí je mírně záporná, jen pro pár společností je možné říct, že bankrot nastal ve skutečnosti později, než je modelováno.

Graf 4.2 Deviantní rezidua pro Coxův model a parametrický model



Zdroj: vlastní zpracování

4.6 Srovnání Coxova a parametrického modelu

V předchozím textu byly pomocí neparametrických metod porovnány funkce přežití napříč jednotlivými sektory a provedeny odhady dvou různých modelů pro sektor doprava, které zahrnují vysvětlující proměnné. Na závěr je provedeno srovnání výsledného semiparametrického modelu a parametrického modelu. Cílem Coxova semiparametrického modelu je vyčíslit vliv jednotlivých veličin na dobu přežití. U parametrického modelu je mimo vyčíslení vlivu nezávislých proměnných také možné určit základní hazardní funkci, za předpokladu, že je určeno rozdělení pravděpodobnosti, kterým se řídí doba přežití jednotlivých subjektů, což je v této práci Weibullovo rozdělení. Do výsledného modelu bylo vybráno šest ukazatelů, stejných pro oba modely. Výsledné hodnoty hazardního poměru pro tyto proměnné a rozdíl mezi nimi je zachycen v tabulce (Tabulka 4.19).

Rozdíly mezi hazardními poměry pro jednotlivé modely nejsou výrazné. Důležitou informací je fakt, že u odhadnutých koeficientů se neliší znaménko a tak hazardní poměr pro každý ukazatel je vždy buď menší než jedna, nebo větší než jedna. Nejvyšší rozdíl je u ukazatele podílu vlastního kapitálu na aktivech a celkové zadluženosti, kdy podle parametrického modelu nárůst těchto ukazatelů vede k vyššímu riziku bankrotu než podle Coxova modelu. Při jednotkovém zvýšení těchto proměnných je podle parametrického modelu vyšší pravděpodobnost bankrotu o 2,1 %, respektive o 2,2 % než podle Coxova modelu. Rovněž u přirozeného logaritmu celkových aktiv klade na proměnnou větší důraz parametrický model. Zvýšení proměnné o jednotku podle parametrického modelu vede ke snížení rizika výskytu koncové události o 1,4 % více než podle Coxova modelu. U stupně krytí stálých aktiv už je rozdíl nepatrný a u doby obraty závazků a rentability nákladů zanedbatelný. Na základě těchto výsledků je tak možné říct, že nezávislé veličiny mají podle parametrického modelu vyšší vliv na riziko bankrotu, než podle Coxova semiparametrického modelu. Obecně by přitom parametrické modely měly dosahovat přesnějších výsledků kvůli použití klasické metody maximální věrohodnosti, ovšem za předpokladu, že je správně určeno pravděpodobnostní rozdělení náhodné veličiny.

Tabulka 4.19 Srovnání hazardních poměrů pro jednotlivé modely

Proměnná	ln Celkových aktiv	Podíl VK na aktivech	Celková zadluženost	Stupeň krytí stálých aktiv	DO závazků	ROC
Hazardní poměr Cox	0,75019	1,09050	1,09742	0,98696	1,000003	1,00007
Hazardní poměr Weibull	0,73592	1,11185	1,11892	0,98858	1,000003	1,00007
Rozdíl	0,01427	-0,02135	-0,02151	-0,00162	0,00000	0,00000

Zdroj: vlastní zpracování

Pro grafické zobrazení výsledků je využito modelových příkladů, kdy jsou zobrazeny funkce přežití a kumulativní hazardní funkce pro určité percentily hodnot. V tabulce (Tabulka 4.20) je vypsán 5%, 25%, 75% a 95% percentil a medián hodnot pro každou proměnnou.

Tabulka 4.20 Percentil hodnot vybraných proměnných

Proměnná/Percentil	5%	25%	50%	75%	95%
ln Celkových aktiv	7,326	8,69	9,789	10,97	12,96
Podíl VK na aktivech	-0,15	0,176	0,375	0,622	0,8877
Celková zadluženost	0,107	0,368	0,613	0,807	1,133
Stupeň krytí stálých aktiv	-0,142	0,399	0,746	2,167	24,714
DO závazků	19,137	43,203	71,294	125,48	426,07
ROC	-0,513	-0,024	0,034	0,16	0,734

Zdroj: vlastní zpracování

Pro namodelování funkce přežití je vybrán 25% a 75% percentil a medián hodnot. Přitom je třeba vzít v úvahu, že koeficienty pro přirozený logaritmus celkových aktiv a stupeň krytí stálých aktiv jsou záporné a ostatní koeficienty jsou kladné. To znamená, že když je sestavován model pro ilustrativní firmu, jejichž hodnoty ukazatelů jsou nejhorší na hranici 25 %, tak je pro logaritmus aktiv a stupeň krytí stálých závazků převzata hodnota 25% percentilu, protože snížení ukazatele vede ke zvýšení pravděpodobnosti bankrotu. Oproti tomu pro zbylé ukazatele je brána hodnota 75% percentilu, neboť čím vyšší hodnota těchto proměnných, tím horší to je pro subjekt. Hodnoty, které obsahuje modelová firma 1, jsou pro přehlednost uvedeny v tabulce (Tabulka 4.21). V této tabulce jsou také uvedeny hodnoty pro modelovou firmu 3, která má nejlepší hodnoty ukazatelů na hranici 25 %. Pro proměnné přirozený logaritmus aktiv a stupeň krytí stálých aktiv je tedy brána hodnota 75% percentilu, jelikož čím vyšší hodnotu mají tyto ukazatele, tím nižší pravděpodobnost bankrotu. Pro zbylé ukazatele je využita hodnota na 25% hranici hodnot.

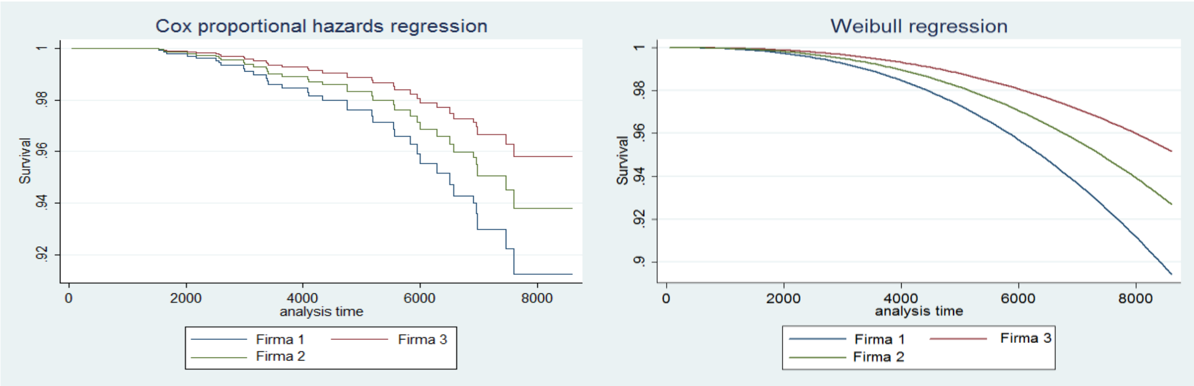
Tabulka 4.21 Hodnoty proměnných pro firmu 1 a firmu 3

	Firma 1					
Proměnná	ln Celkových aktiv	Podíl VK na aktivech	Celková zadluženost	Stupeň krytí stálých aktiv	DO závazků	ROC
Hodnota	8,69	0,622	0,807	0,399	125,48	0,16
	Firma 3					
Proměnná	ln Celkových aktiv	Podíl VK na aktivech	Celková zadluženost	Stupeň krytí stálých aktiv	DO závazků	ROC
Hodnota	10,97	0,176	0,368	2,167	43,203	-0,024

Zdroj: vlastní zpracování

Na obrázku (Obrázek 4.6) jsou znázorněny funkce přežití pro modelové firmy 1, 2 a 3. Firmy 1 a 3 jsou popsány v předcházejícím odstavci, firma 2 obsahuje medián hodnot jednotlivých proměnných. V levé části jsou funkce přežití pro Coxův model (3.15) , kdy je charakteristická schodovitost funkce, v pravé části jsou funkce přežití vyhlazené, jelikož se jedná o parametrický model (3.32), kdy je předpoklad o Weibullovu rozdělení pravděpodobnosti doby přežití subjektu. Z obou grafů je patrné, že přímky jsou paralelní, a to odpovídá tomu, že se jedná o modely, jejichž předpokladem je proporcionalita hazardu v čase. Očekávání odpovídá i pořadí křivek, jelikož nejrychleji klesá pravděpodobnost přežití pro firmu 1, ve které jsou obsaženy nejhorší hodnoty na hranici 25 %. Uprostřed leží křivka přežití pro firmu 2, která obsahuje medián hodnot a nejvyšší pravděpodobnost přežití má firma 3.

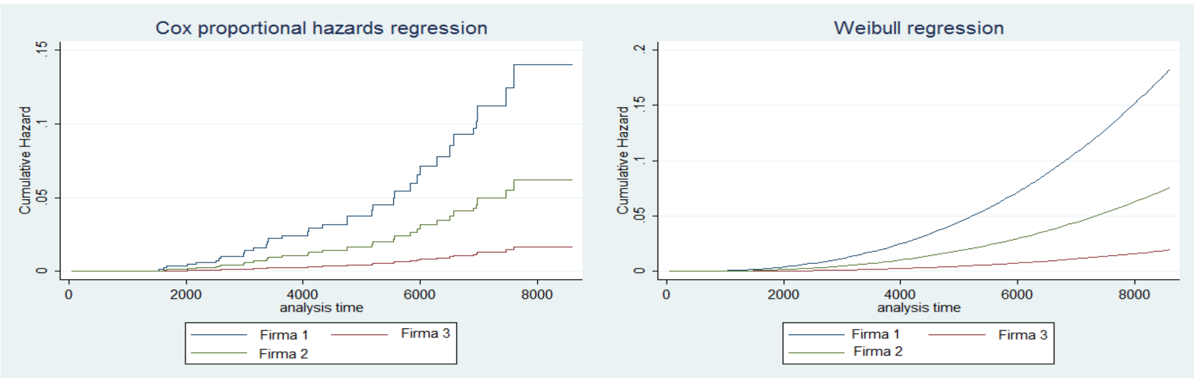
Obrázek 4.6 Funkce přežití pro ilustrativní firmy podle Coxova a parametrického modelu



Zdroj: vlastní zpracování

Podobné výsledky jsou patrné i z následujícího obrázku (Obrázek 4.7), kde jsou zobrazeny kumulativní hazardní funkce pro modelové firmy 1 až 3, v levé části podle Coxova modelu (4.6) a v pravé části podle parametrického modelu (4.11). Firma 2 opět obsahuje medián hodnot, firma 1 ale nyní má nejhorší hodnoty na 5% hranici a firma 3 naopak nejlepší hodnoty na 95% hranici (Tabulka 4.20). Jsou tak více zvýrazněny rozdíly mezi jednotlivými křivkami v rámci grafů, kdy ke konci sledované období je pravděpodobnost bankrotu pro firmu 1 vyšší o více jak 10 %. Obdobně jako u funkce přežití jsou si i kumulativní hazardní funkce podle Coxova a parametrického modelu celkem podobné, až na to, že v levé části jsou funkce schodovité a v pravé vyhlazené.

Obrázek 4.7 Kumulativní hazardní funkce pro ilustrativní firmy podle Coxova a parametrického modelu



Zdroj: vlastní zpracování

5 Závěr

Cílem této práce bylo posouzení doby přežití českých společností od jejich vzniku do bankrotu těchto podniků. Nejprve byla zkoumána data 16513 českých společností z 9 různých odvětví, z nichž u 1504 společností došlo k bankrotu mezi roky 2005 až 2015, za účelem porovnání doby přežití napříč jednotlivými sektory. Využitím Kaplan-Meierova odhadu funkce přežití a Nelson-Aalenova odhadu kumulativní hazardní funkce bylo zjištěno, že největší pravděpodobnost úpadku mají podniky působící v odvětví stavebnictví, ubytování a doprava, kde pravděpodobnost doby přežití v čase výrazně klesá. Oproti tomu společnosti působící v zemědělství, zdravotnictví nebo zabývající se zpracováním vody a odpadu mají nízkou pravděpodobnost výskytu bankrotu a pravděpodobnost doby přežití klesá v čase jen minimálně.

Vzhledem k rozdílnosti jednotlivých sektorů bylo sestavení Coxova semiparametrického modelu a parametrického modelu provedeno pouze pro podniky z jednoho odvětví. Vybrán byl sektor doprava, a to z důvodu velkého počtu podniků a vysokého počtu úpadků, navíc se toto odvětví v neparametrické analýze jevilo jako specifické, neboť funkce přežití pro tento sektor nebyla shodná s žádnou funkcí přežití pro ostatní odvětví. Bylo tedy analyzováno 1229 českých společností z tohoto sektoru, z nichž u 232 došlo k bankrotu mezi roky 2005 až 2015. K dispozici bylo celkem 24 finančních ukazatelů pro každou společnost a tyto informace sloužily jako vstupní údaje pro sestavení Coxova modelu a parametrického modelu za předpokladu, že doba přežití odpovídá Weibullovu rozdělení pravděpodobnosti. Nejprve bylo nutné provést užší výběr proměnných a po důkladné analýze bylo vybráno 6 proměnných, které ovlivňují dobu přežití podniku. Tyto proměnné byly pro oba sestavované modely stejné a jedná se o přirozený logaritmus aktiv, stupeň krytí stálých aktiv, celkovou zadluženost, podíl vlastního kapitálu na aktivech, dobu obratu závazků a rentabilitu nákladů.

U přirozeného logaritmu aktiv a stupně krytí stálých aktiv byl odhadnutý koeficient záporný, to znamená, že rostoucí hodnota tohoto ukazatele vede ke snížení pravděpodobnosti bankrotu. Koeficient pro celkovou zadluženost a podíl vlastního kapitálu na aktivech vyšel kladný. Vyšší hodnota těchto proměnných tak vede ke zvýšení rizika úpadku firmy, což je v souladu s očekáváním. Koeficienty pro dobu obratu závazků a rentabilitu nákladů jsou téměř nulové a nemají tak prakticky žádný vliv na dobu přežití jednotlivých společností. Součástí parametrického modelu je také odhadnutí parametrů měřítka a tvaru, díky čemuž mohla být odhadnuta základní hazardní funkce, která odpovídá Weibullovu rozdělení pravděpodobnosti. Hodnota parametru tvaru p je větší než jedna, základní hazardní funkce, která je pro všechny firmy stejná, je tak rostoucí, riziko bankrotu v čase roste.

Výsledky Coxova a parametrického modelu se příliš neliší. U obou modelů byly do konečného modelu vybrány stejné proměnné, hodnoty koeficientů se sice mírně odlišují, ale znaménko u jednotlivých koeficientů je vždy stejné. Rovněž využitím Cox-Snellových, martingalových a deviantních reziduí nebyl zjištěn významný rozdíl v adekvátnosti jednotlivých modelů, což znamená, že není možné určit, který model je vhodnější. Obecně by parametrický model měl být přesnější, avšak je výpočetně náročnější a vzhledem k tomu, že výsledky se v tomto případě liší jen minimálně, tak z hlediska praktického využití je dostačující použití Coxova semiparametrického modelu.

Práci je možné dále rozšířit o sestavení parametrického modelu s jiným než Weibullovým rozdělením pravděpodobnosti, například s exponenciálním, log-logistickým, nebo zobecněným gamma rozdělením. Pro tato rozdělení, ale i pro Weibullovo rozdělení, lze rovněž sestavit AFT model, kdy už není nutná podmínka o proporcionalitě hazardu a základní hazardní křivka je tedy pro každý subjekt jiná. Možností by bylo také využití jiných metod pro výběr proměnných do konečného modelu, neboť by to mohlo vést k jiným výsledkům.

Seznam použité literatury

Monografie

[1] CLEVES, Mario Alberto. *An introduction to survival analysis using Stata*. 3rd ed. College Station, Tex.: Stata Press, 2010, xxviii, ISBN 978-1-59718-074-0.

[2] COLLETT, D. *Modelling survival data in medical research*. 2nd ed. Boca Raton: Chapman & Hall/CRC, c2003. Text in statistical science. ISBN 1-58488-325-1.

[3] COX, D. R. *Partial likelihood*. Biometrika 1975, vol. 62

[4] COX, D.R. *Regression Models and Life-Tables (with discussion)*. Journal of the Royal Statistical Society 1972, Series B, vol. 34

[5] DLUHOŠOVÁ, Dana. *Finanční řízení a rozhodování podniku: analýza, investování, oceňování, riziko, flexibilita*. 3., rozš. vyd. Praha: Ekopress, 2010, ISBN 978-80-86929-68-2.

[6] FOTR, Jiří a Ivan SOUČEK. *Podnikatelský záměr a investiční rozhodování*. 1. vyd. Praha: Grada Publishing, 2005. Expert (Grada). ISBN 80-247-0939-2.

[7] GILL, R. *Censoring and stochastic integrals*. Amsterdam: Mathematisch Centrum, 1980. ISBN 9061961971.

[8] HAIR, Joseph F. *Multivariate data analysis*. 7th ed. Harlow: Pearson, c2014. Pearson new international edition. ISBN 978-1-292-02190-4.

[9] HANČLOVÁ, Jana. *Ekonometrické modelování: klasické přístupy s aplikacemi*. 1. vyd. Praha: Professional Publishing, 2012. ISBN 978-80-7431-088-1

[10] HOSMER, David W and Stanley LEMESHOW. *Applied logistic regression*. 2nd ed. New York: John Wiley & Sons, c2000. Wiley series in probability and statistics. ISBN 0-471-35632-8.

[11] HOSMER, D., W. S. LEMESHOW and S. MAY. *Applied survival analysis: regression modeling of time-to-event data*. 2nd ed. Hoboken: Wiley-Interscience, c2008. xiii, ISBN 978-0-471-75499-2.

[12] NELSON, Wayne. *Applied life data analysis*. New York: Wiley, c1982. ISBN 0471094587.

[13] ROYSTON, Patrick and Paul C LAMBERT. *Flexible parametric survival analysis using stata: beyond the Cox model*. College Station, TX: Stata Press, 2011. ISBN 978-1-59718-079-5.

[14] SMEJKAL, Vladimír a Karel RAIS. *Řízení rizik ve firmách a jiných organizacích*. 2., aktualiz. a rozš. vyd. Praha: Grada, 2006. Expert (Grada). ISBN 80-247-1667-4.

[15] VOCHOZKA, Marek. *Metody komplexního hodnocení podniku*. 1. vyd. Praha: Grada, 2011. Finanční řízení. ISBN 978-80-247-3647-1.

Elektronické dokumenty a ostatní

[1] Bisnode. *MagnusWeb* [online]. [cit. 2015-06-08]. Dostupné z: : http://magnusweb.bisnode.cz

[2] IBRAHIM, Joseph. *Survival analysis: Introduction* [online]. In: . s. 491 [cit. 2016-04-10]. Dostupné z: http://www.amstat.org/chapters/northeasternillinois/pastevents/presentations/summer05_Ibrahim_J.pdf.

[3] Idre UCLA. *Survival analysis with Stata* [online]. [cit. 2015-08-25]. Dostupné z: http://www.ats.ucla.edu/stat/stata/seminars/stata_survival/.

[4] Janurová, Kateřina. *Vyspělé statistické metody pro efektivní vyhodnocení biomedicínských dat.* Ostrava, 2014. Disertační práce. Vysoká škola báňská – Technická univerzita Ostrava, Fakulta elektrotechniky a informatiky, Katedra kybernetiky a biomedicínského inženýrství.

[5] Matematická biologie. *P-hodnota a její interpretace* [online]. [cit. 2016-01-06]. Dostupné z: http://portal.matematickabiologie.cz/index.php?pg=aplikovana-analyza-klinickych-a-biologickych-dat--biostatistika-pro-matematickou-biologii--uvod-do-testovani-hypotez--p-hodnota-a-jeji-interpretace.

[6] NCSS.com. *Distribution (Weibull) Fitting* [online]. [cit. 2016-01-10]. Dostupné z: https://www.researchgate.net/file.PostFileLoader.html.

[7] Stata manual. *Stcox postestimation* [online]. [cit. 2016-04-10]. Dostupné z: http://www.stata.com/manuals13/ststcoxpostestimation.pdf

[8] Stata manual. *Streg postestimation* [online]. [cit. 2015-12-17]. Dostupné z: http://www.stata.com/manuals13/ststregpostestimation.pdf.

[9] Wikiskripta. *Střední chyba průměru* [online]. [cit. 2016-04-10]. Dostupné z: http://www.wikiskripta.eu/index.php/St%C5%99edn%C3%AD_chyba_pr%C5%AFm%C4%9Bru.

[10] Zákon č. 182/2006 Sb., o úpadku a způsobech jeho řešení. In: Sbírka zákonů České republiky. 2006. Dostupný také z: http://business.center.cz/business/pravo/zakony/insolvencni/.

Seznam zkratek

AFT	Model zrychleného času
ČPK	Čistý pracovní kapitál
DCK	Dlouhodobý cizí kapitál
EAT	Zisk po zdanění
EBIT	Zisk před zdaněním a odpisy
HR	Hazardní poměr
IT	Informační technologie
OA	Oběžná aktiva
PH	Proporcionální hazard
ROA	Rentabilita aktiv
ROC	Rentabilita nákladů
ROCE	Rentabilita dlouhodobých zdrojů
ROE	Rentabilita vlastního kapitálu
ROS	Rentabilita tržeb
T	Tržby
VK	Vlastní kapitál

Prohlášení o využití výsledků diplomové (bakalářské) práce

Prohlašuji, že

- jsem byl seznámen s tím, že na mou diplomovou práci se plně vztahuje zákon č. 121/2000 Sb. – autorský zákon, zejména § 35 – užití díla v rámci občanských a náboženských obřadů, v rámci školních představení a užití díla školního a § 60 – školní dílo;
- beru na vědomí, že Vysoká škola báňská – Technická univerzita Ostrava (dále jen VŠB-TUO) má právo nevýdělečně, ke své vnitřní potřebě, diplomovou práci užít (§ 35 odst. 3);
- souhlasím s tím, že diplomová práce bude v elektronické podobě archivována v Ústřední knihovně VŠB-TUO a jeden výtisk bude uložen u vedoucího diplomové (bakalářské) práce. Souhlasím s tím, že bibliografické údaje o diplomové práci budou zveřejněny v informačním systému VŠB-TUO;
- bylo sjednáno, že s VŠB-TUO, v případě zájmu z její strany, uzavřu licenční smlouvu s oprávněním užít dílo v rozsahu § 12 odst. 4 autorského zákona;
- bylo sjednáno, že užít své dílo, diplomovou práci, nebo poskytnout licenci k jejímu využití mohu jen se souhlasem VŠB-TUO, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly VŠB-TUO na vytvoření díla vynaloženy (až do jejich skutečné výše).

V Ostravě dne 21. 4. 2016

Klimek

....

Bc. Vojtěch Klimek

Seznam příloh

Příloha č. 1: Detailní výsledky logrank testu

Sektor	Skutečný počet bankrotů	Očekávaný počet bankrotů			
Administrativa	127	123,39	chi2(1)	=	0,39
Kultura	43	46,61	Pr>chi2	=	0,5343
Administrativa	127	209,18	chi2(1)	=	77,59
Doprava	232	149,82	Pr>chi2	=	0
Administrativa	127	94,19	chi2(1)	=	19,8
IT	96	128,81	Pr>chi2	=	0
Administrativa	127	91,38	chi2(1)	=	23,94
Zemědělství	95	130,62	Pr>chi2	=	0
Administrativa	127	113,07	chi2(1)	=	5,88
Zpracování	33	46,93	Pr>chi2	=	0,0153
Administrativa	127	207,44	chi2(1)	=	42,45
Stavebnictví	669	588,56	Pr>chi2	=	0
Administrativa	127	179,46	chi2(1)	=	34,62
Ubytování	196	143,54	Pr>chi2	=	0
Administrativa	127	101,06	chi2(1)	=	24,54
Zdravotnictví	12	37,94	Pr>chi2	=	0
Doprava	232	180,08	chi2(1)	=	43,56
Kultura	43	94,92	Pr>chi2	=	0
IT	96	109,26	chi2(1)	=	7,54
Kultura	43	29,74	Pr>chi2	=	0,006
Kultura	43	27,68	chi2(1)	=	10,77
Zemědělství	95	110,32	Pr>chi2	=	0,001
Kultura	43	35,88	chi2(1)	=	2,70
Zpracování	33	40,12	Pr>chi2	=	0,1004
Kultura	43	83,08	chi2(1)	=	22
Stavebnictví	669	628,92	Pr>chi2	=	0
Kultura	43	76,65	chi2(1)	=	21,78
Ubytování	196	162,35	Pr>chi2	=	0
Kultura	43	26,83	chi2(1)	=	19,3
Zdravotnictví	12	28,17	Pr>chi2	=	0
Doprava	232	113	chi2(1)	=	191,33
IT	96	215	Pr>chi2	=	0
Doprava	232	111,23	chi2(1)	=	199,7
Zemědělství	95	215,77	Pr>chi2	=	0
Doprava	232	168,31	chi2(1)	=	66,12
Zpracování	33	96,69	Pr>chi2	=	0
Doprava	232	182,66	chi2(1)	=	16,73
Stavebnictví	669	718,34	Pr>chi2	=	0
Doprava	232	201,75	chi2(1)	=	8,68
Ubytování	196	226,25	Pr>chi2	=	0,0032

Doprava	232	160,23	chi2(1)	=	93,88
Zdravotnictví	12	83,77	Pr>chi2	=	0
IT	96	92,92	chi2(1)	=	0,2
Zemědělství	95	98,08	Pr>chi2	=	0,6544
IT	96	98,82	chi2(1)	=	0,35
Zpracování	33	30,18	Pr>chi2	=	0,5569
IT	96	248,79	chi2(1)	=	139,6
Stavebnictví	669	516,21	Pr>chi2	=	0
IT	96	184,14	chi2(1)	=	115,06
Ubytování	196	107,86	Pr>chi2	=	0
IT	96	84,67	chi2(1)	=	7,70
Zdravotnictví	12	23,33	Pr>chi2	=	0,0079
Zemědělství	95	99	chi2(1)	=	0,71
Zpracování	33	29	Pr>chi2	=	0,3979
Stavebnictví	669	509,99	chi2(1)	=	149,55
Zemědělství	95	254,01	Pr>chi2	=	0
Ubytování	196	106,42	chi2(1)	=	122,45
Zemědělství	95	184,58	Pr>chi2	=	0
Zdravotnictví	12	22,21	chi2(1)	=	5,94
Zemědělství	95	84,79	Pr>chi2	=	0,0148
Stavebnictví	669	612,41	chi2(1)	=	41,01
Zpracování	33	89,59	Pr>chi2	=	0
Ubytování	196	152,02	chi2(1)	=	38,44
Zpracování	33	76,98	Pr>chi2	=	0
Zdravotnictví	12	21,46	chi2(1)	=	7,98
Zpracování	33	23,54	Pr>chi2	=	0,0047
Stavebnictví	669	674,95	chi2(1)	=	0,24
Ubytování	196	190,05	Pr>chi2	=	0,623
Stavebnictví	669	601,54	chi2(1)	=	64,95
Zdravotnictví	12	79,46	Pr>chi2	=	0
Ubytování	196	142,59	chi2(1)	=	64,84
Zdravotnictví	12	65,41	Pr>chi2	=	0

Příloha č. 2: Detailní výsledky Wilcoxonova testu

Sektor	Skutečný počet bankrotů	Očekávaný počet bankrotů			
Administrativa	127	123,39	chi2(1)	=	1,50
Kultura	43	46,61	Pr>chi2	=	0,221
Administrativa	127	209,18	chi2(1)	=	75,75
Doprava	232	149,82	Pr>chi2	=	0
Administrativa	127	94,19	chi2(1)	=	18,59
IT	96	128,81	Pr>chi2	=	0
Administrativa	127	91,38	chi2(1)	=	26,74
Zemědělství	95	130,62	Pr>chi2	=	0
Administrativa	127	113,07	chi2(1)	=	3,58
Zpracování	33	46,93	Pr>chi2	=	0,0585
Administrativa	127	207,44	chi2(1)	=	37,13
Stavebnictví	669	588,56	Pr>chi2	=	0
Administrativa	127	179,46	chi2(1)	=	29,72
Ubytování	196	143,54	Pr>chi2	=	0
Administrativa	127	101,06	chi2(1)	=	19,27
Zdravotnictví	12	37,94	Pr>chi2	=	0
Doprava	232	180,08	chi2(1)	=	48,4
Kultura	43	94,92	Pr>chi2	=	0
IT	96	109,26	chi2(1)	=	3,63
Kultura	43	29,74	Pr>chi2	=	0,0568
Kultura	43	27,68	chi2(1)	=	7,69
Zemědělství	95	110,32	Pr>chi2	=	0,0056
Kultura	43	35,88	chi2(1)	=	0,4
Zpracování	33	40,12	Pr>chi2	=	0,526
Kultura	43	83,08	chi2(1)	=	24,36
Stavebnictví	669	628,92	Pr>chi2	=	0
Kultura	43	76,65	chi2(1)	=	23,70
Ubytování	196	162,35	Pr>chi2	=	0
Kultura	43	26,83	chi2(1)	=	11,11
Zdravotnictví	12	28,17	Pr>chi2	=	0,0009
Doprava	232	113	chi2(1)	=	183,49
IT	96	215	Pr>chi2	=	0
Doprava	232	111,23	chi2(1)	=	198,34
Zemědělství	95	215,77	Pr>chi2	=	0
Doprava	232	168,31	chi2(1)	=	57,24
Zpracování	33	96,69	Pr>chi2	=	0
Doprava	232	182,66	chi2(1)	=	20,25
Stavebnictví	669	718,34	Pr>chi2	=	0
Doprava	232	201,75	chi2(1)	=	10,15
Ubytování	196	226,25	Pr>chi2	=	0,0014

Doprava	232	160,23	chi2(1)	=	82,9
Zdravotnictví	12	83,77	Pr>chi2	=	0
IT	96	92,92	chi2(1)	=	0,56
Zemědělství	95	98,08	Pr>chi2	=	0,4527
IT	96	98,82	chi2(1)	=	1,23
Zpracování	33	30,18	Pr>chi2	=	0,2681
IT	96	248,79	chi2(1)	=	125,65
Stavebnictví	669	516,21	Pr>chi2	=	0
IT	96	184,14	chi2(1)	=	96,51
Ubytování	196	107,86	Pr>chi2	=	0
IT	96	84,67	chi2(1)	=	4,95
Zdravotnictví	12	23,33	Pr>chi2	=	0,026
Zemědělství	95	99	chi2(1)	=	2,10
Zpracování	33	29	Pr>chi2	=	0,1471
Stavebnictví	669	509,99	chi2(1)	=	138,79
Zemědělství	95	254,01	Pr>chi2	=	0
Ubytování	196	106,42	chi2(1)	=	108,9
Zemědělství	95	184,58	Pr>chi2	=	0
Zdravotnictví	12	22,21	chi2(1)	=	4,54
Zemědělství	95	84,79	Pr>chi2	=	0,0331
Stavebnictví	669	612,41	chi2(1)	=	33
Zpracování	33	89,59	Pr>chi2	=	0
Ubytování	196	152,02	chi2(1)	=	27,2
Zpracování	33	76,98	Pr>chi2	=	0
Zdravotnictví	12	21,46	chi2(1)	=	7,68
Zpracování	33	23,54	Pr>chi2	=	0,0056
Stavebnictví	669	674,95	chi2(1)	=	0,18
Ubytování	196	190,05	Pr>chi2	=	0,6673
Stavebnictví	669	601,54	chi2(1)	=	56,46
Zdravotnictví	12	79,46	Pr>chi2	=	0
Ubytování	196	142,59	chi2(1)	=	48
Zdravotnictví	12	65,41	Pr>chi2	=	0

Příloha č. 3: Základní popisné údaje o nezávislých proměnných

Proměnná	Počet pozorování	Střední hodnota	Směrodatná odchylka	Minimum	Maximum
Celková aktiva	8864	220828	2639850	-673	90300000
ln Celkových aktiv	8855	14,65819	447,5769	-0,9002	42127
ROA	8857	4,734906	446,6991	-214,52	42039
ROE	8850	33,60251	1173,634	-2671,4	45658
ROS	8758	4,81301	450,1749	-267,818	42125
ROCE	8849	33,20869	1184,763	-1147,33	42221
ROC	8797	34,23603	1204,632	-328	46023
Podíl VK na OA	8857	0,266382	6,207639	-320	329
Celková zadluženost	8857	0,764092	5,143348	-2,05498	321
Zadluženost VK	8849	971,0347	40548,51	-1717	2736490
Stupeň krytí stálých aktiv	8505	1486,606	51387,35	-363,179	2730524
Finanční páka	8849	551,1832	28385,38	-1722	2391610
Úrokové krytí	5885	5559,877	97167,02	-23845	2730249
Obrátka aktiv	8853	201,5622	12775,77	-1,26117	977057
DO aktiv	8754	670,4451	16096,74	-7056,79	1388931
DO zásob	8754	88,34608	5469,549	-627,654	498528
DO pohledávek	8754	224,5274	6641,636	-188,932	593897,1
DO závazků	8757	706,2141	28729,71	-5691,43	2360572
Běžná likvidita	8819	916,0387	38705,43	-489	2707210
Pohotová likvidita	8819	970,8902	38218,14	-489	2707210
Okamžitá likvidita	8820	430,0217	22890,14	-473	1817418
Pohledávky na OA	8849	0,672163	0,3994072	-8,84783	21,66667
Zásoby na OA	8849	343,3457	32211,32	-1742	3030083
Čistý pracovní kapitál	8864	10202,88	331024,1	-1,84E+07	2536032